

Figure 1

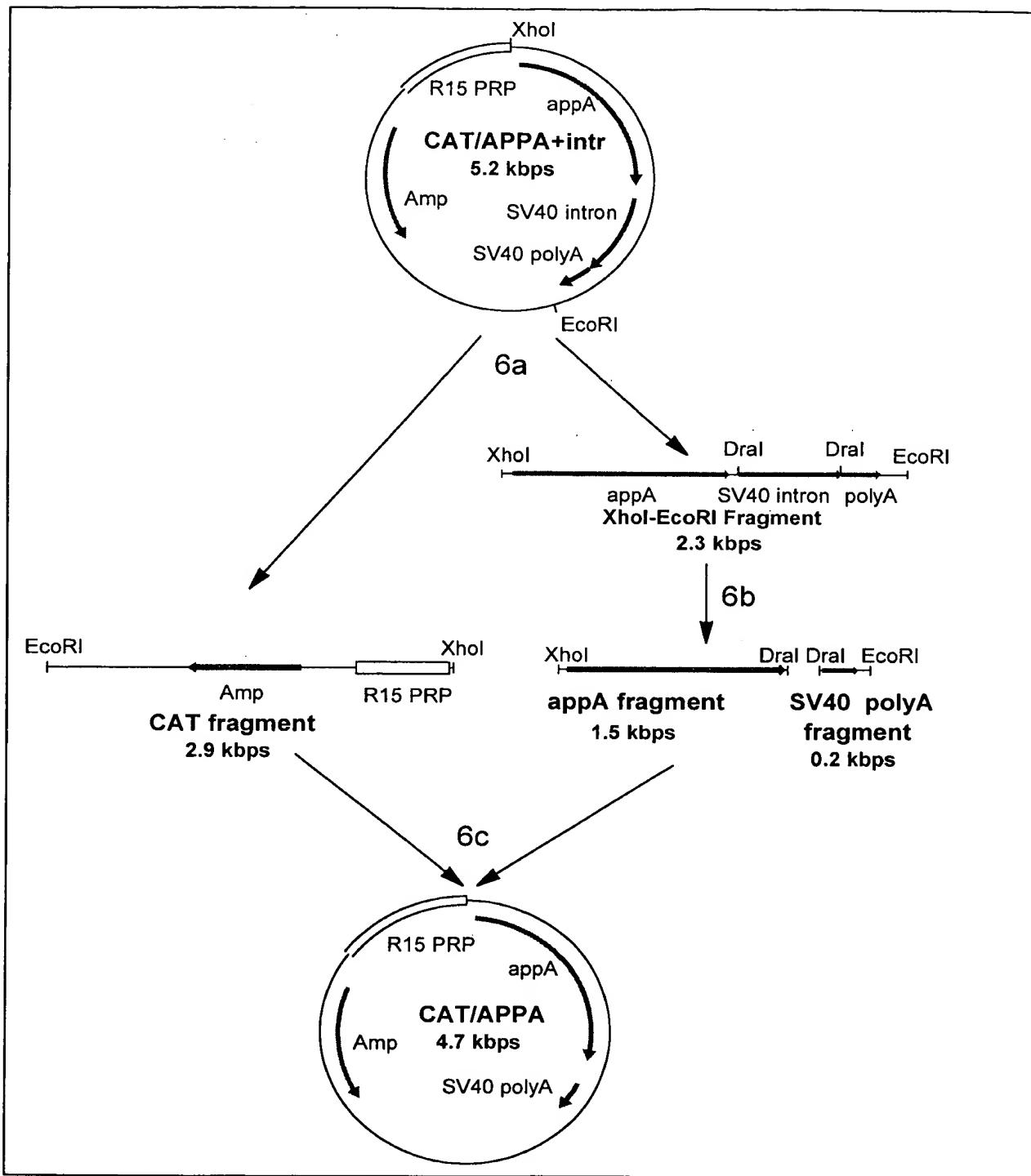


Figure 1A

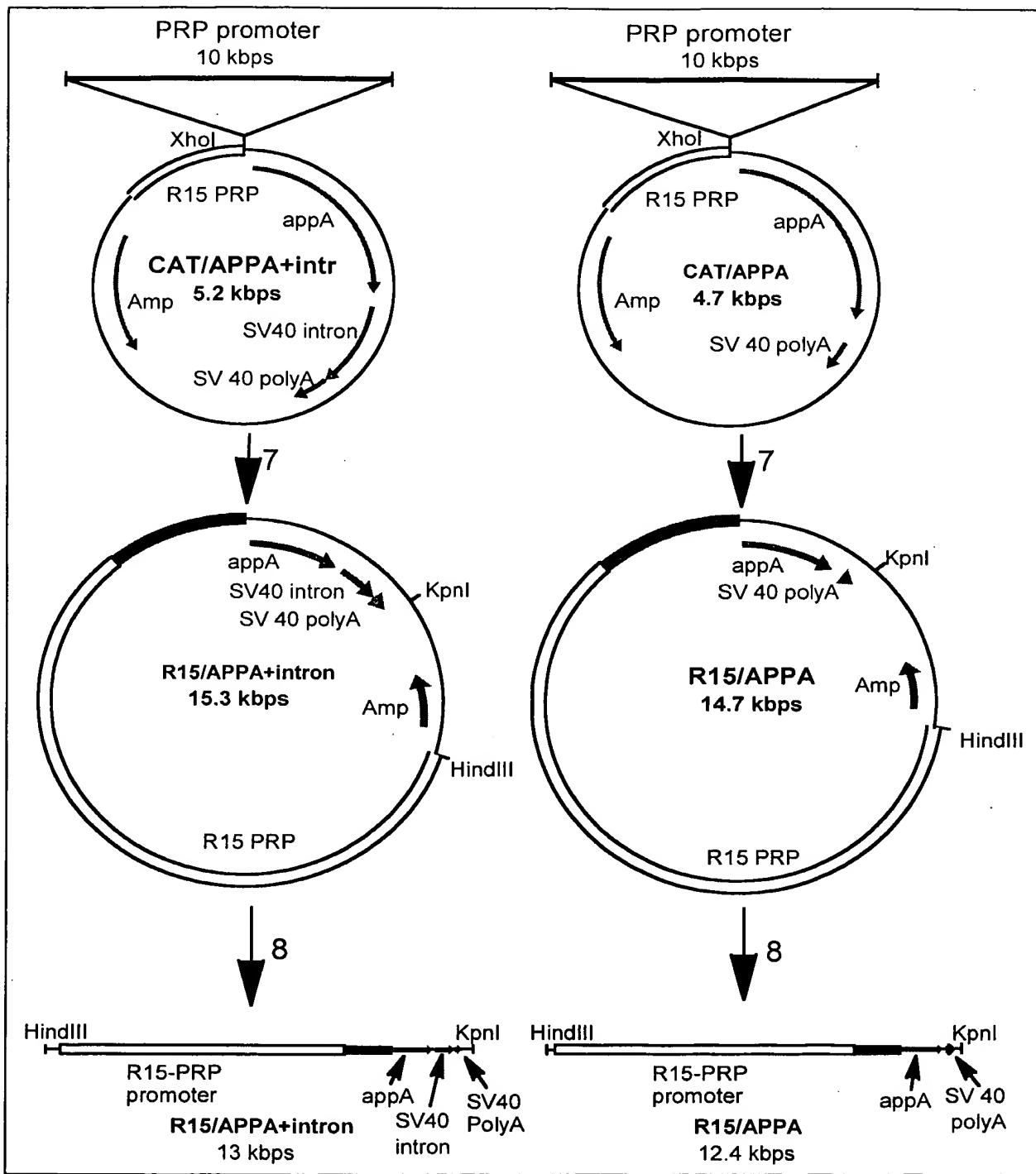


Figure 1B

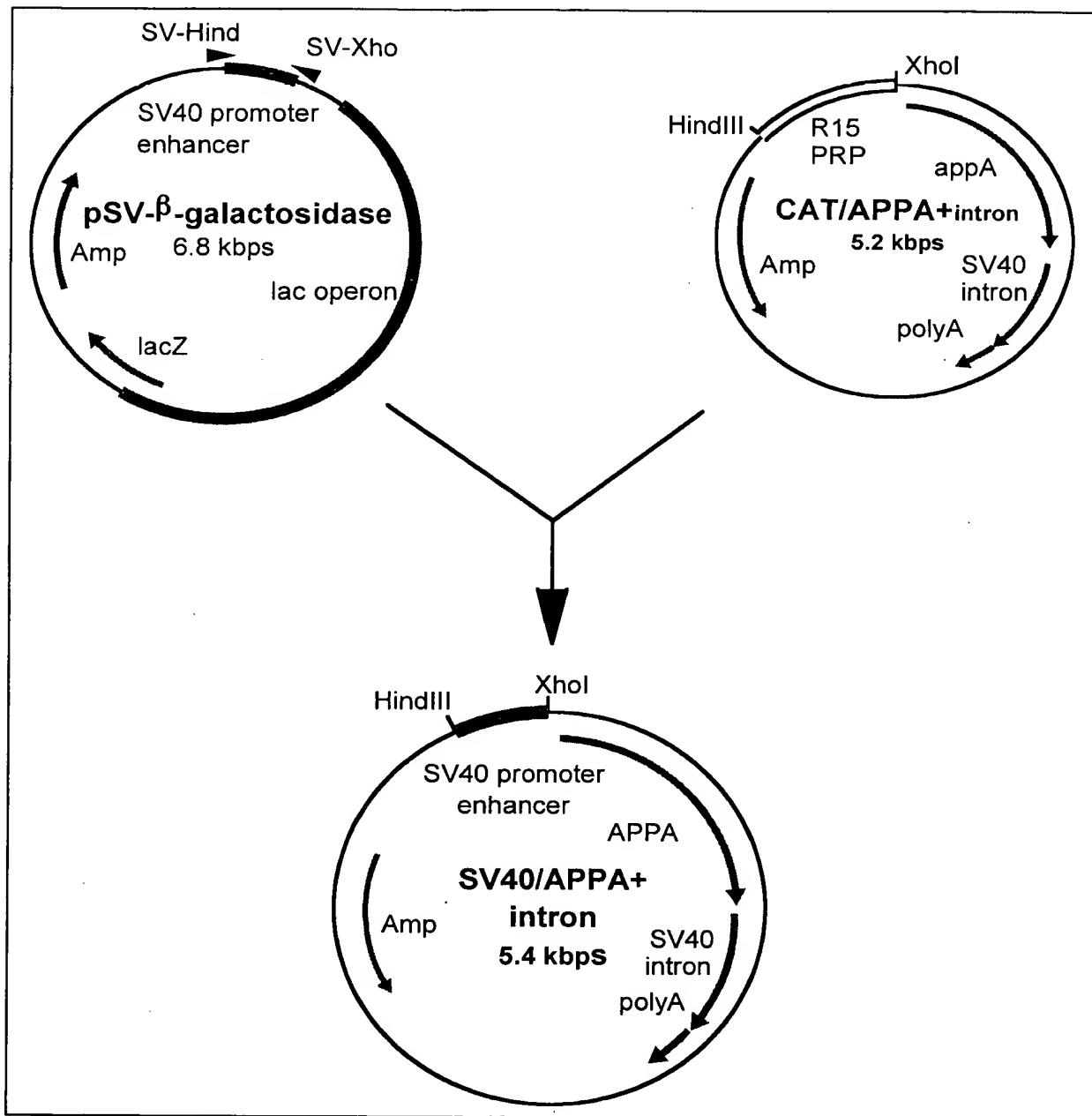


Figure 2

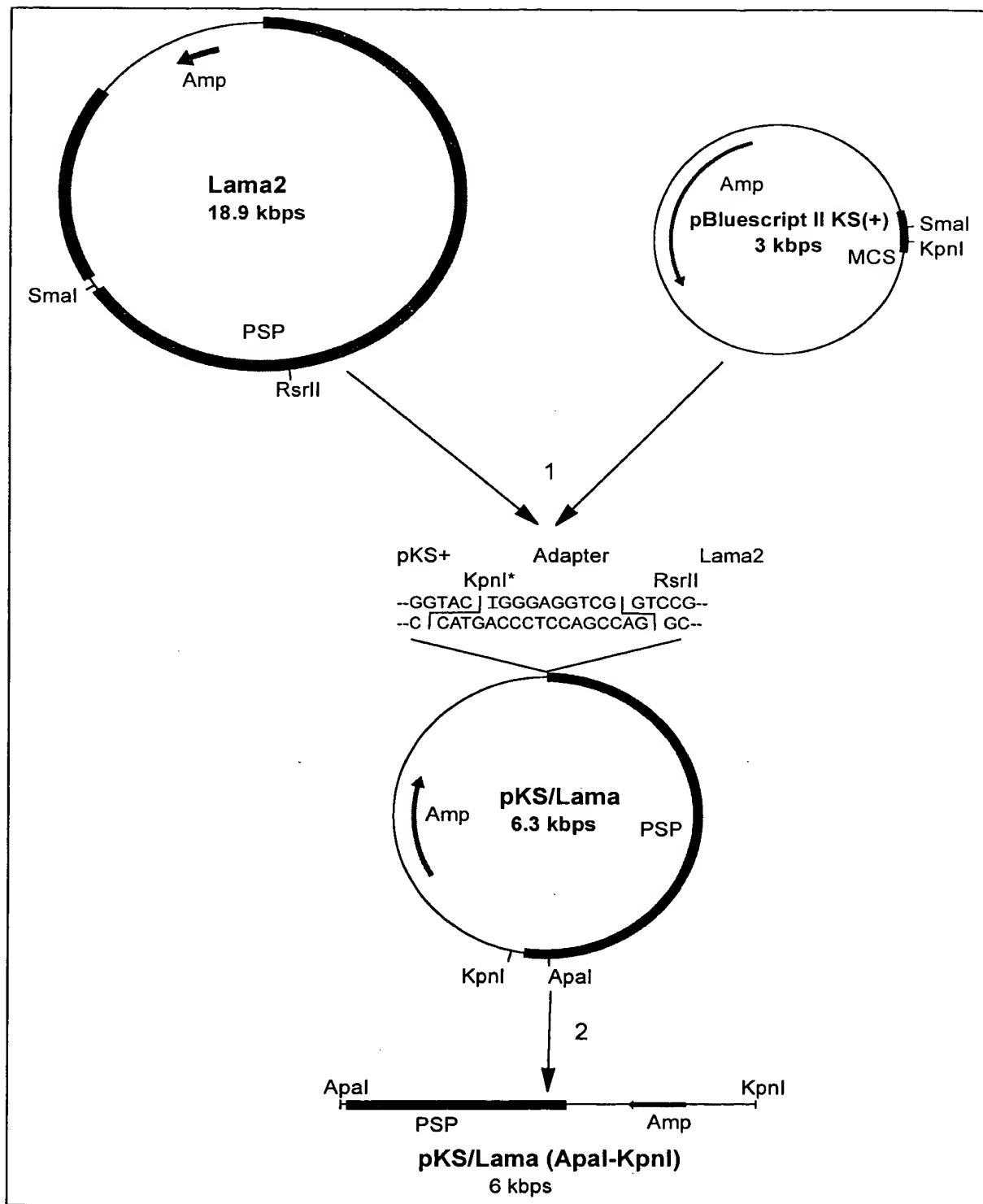


Figure 3

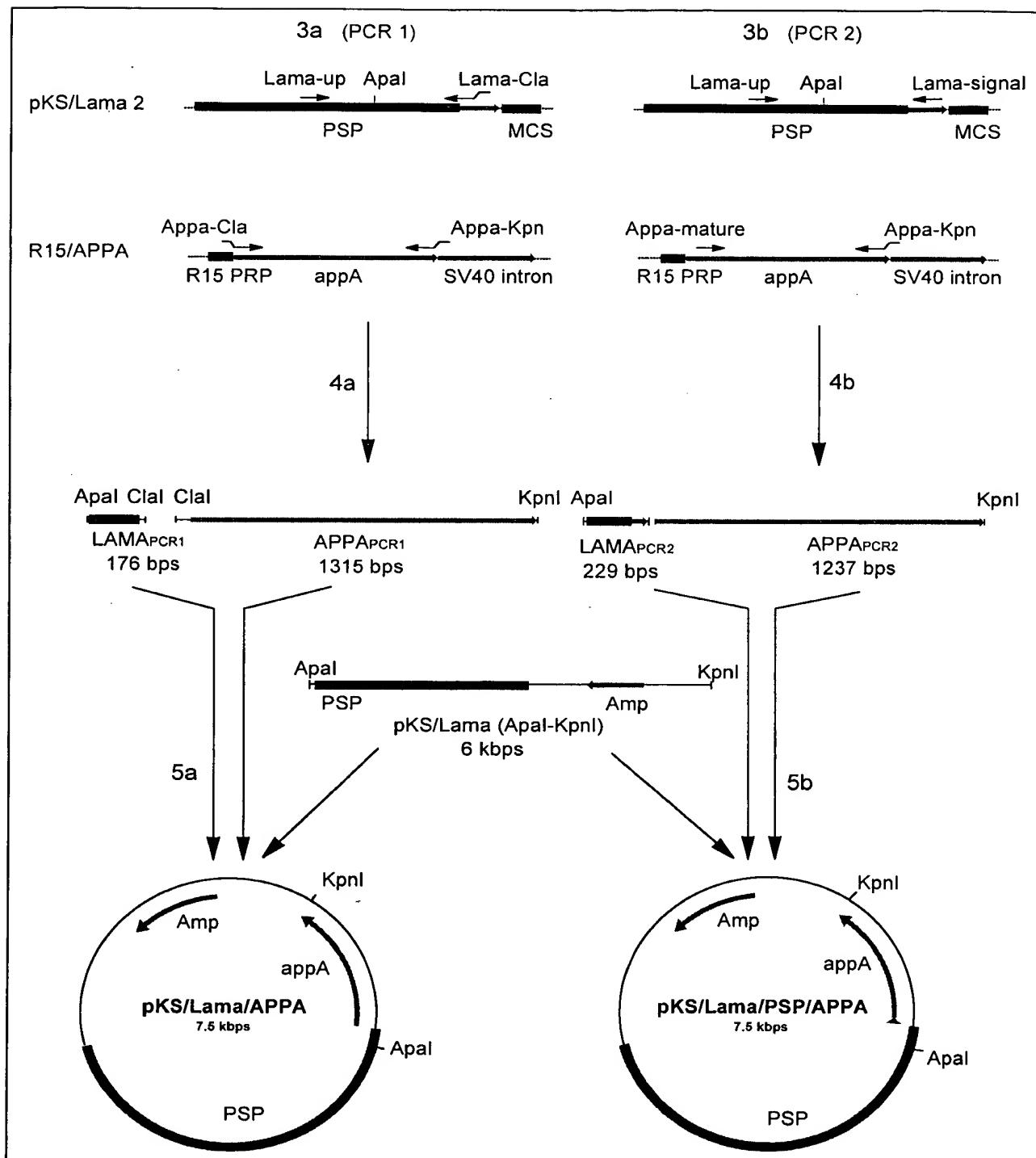


Figure 3A

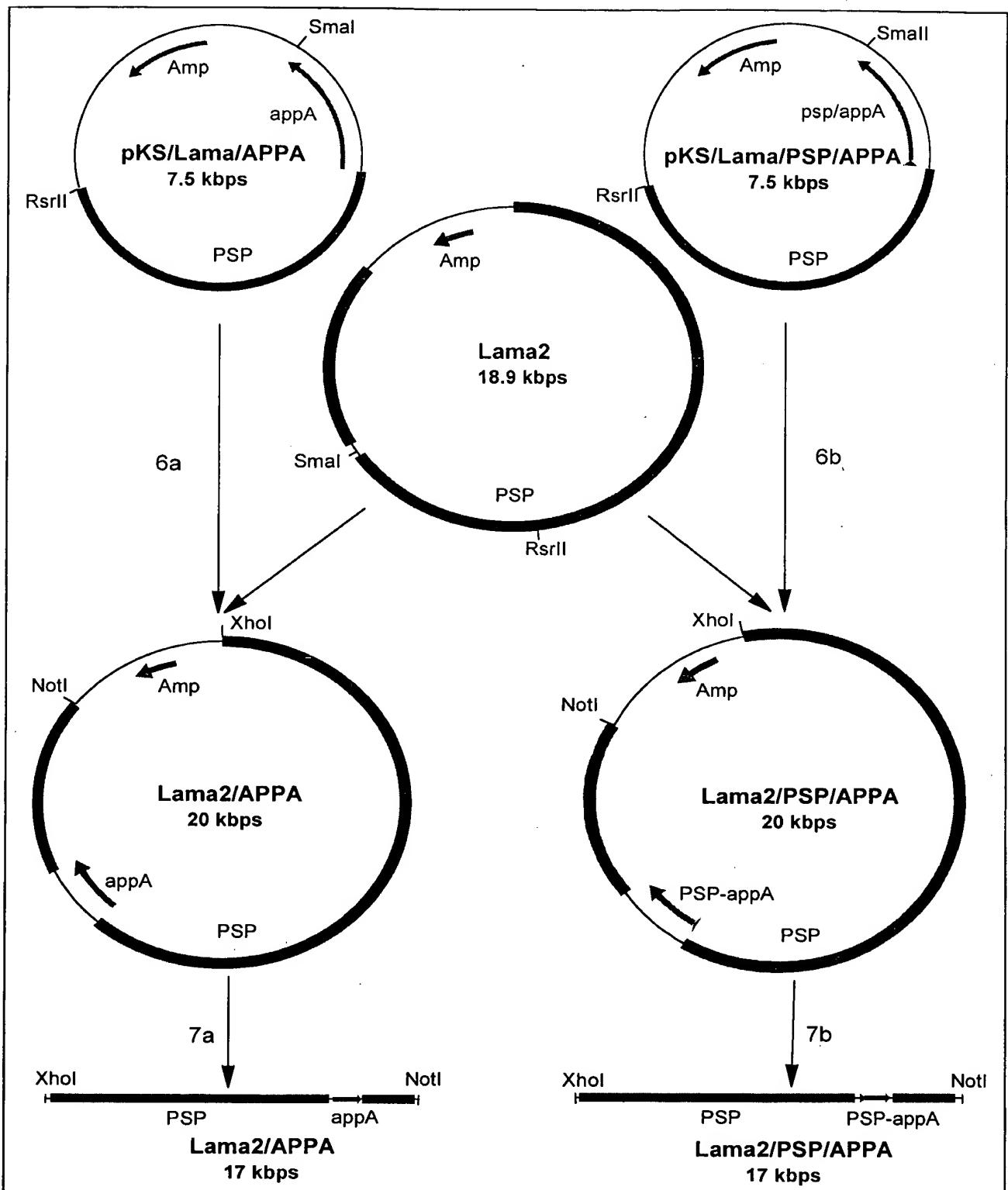


Figure 3B

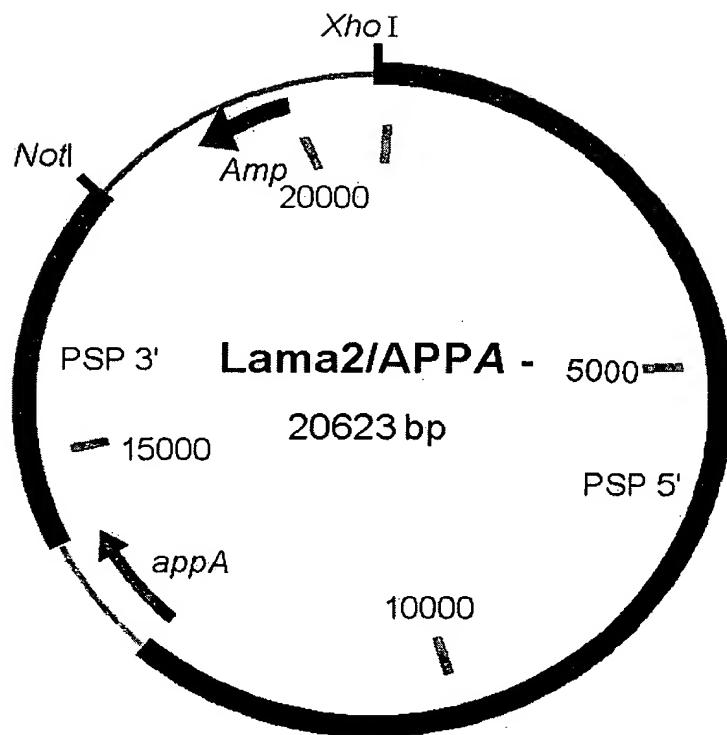


Figure 4. Schematic diagram of the Lama2/APPA construct.

Figure 5. The nucleic acid sequence of the Lama2/APPA plasmid (SEQ ID NO: 1)

LOCUS Lama-appA 20623 bp DNA CIRCULAR SYN 17-JAN-2000
DEFINITION Lama 2/APPA transgenic construct
ACCESSION Lama 2-appA,
KEYWORDS parotid secretory protein; acid glucose-1-phosphatase; appA gene; periplasmic phosphoanhydride phosphohydrolase; artificial sequence; cloning vector
REFERENCE 1 (bases 1 to 20623)
AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
JOURNAL Unpublished.
FEATURES
DEFINITION M. musculus Psp gene for parotid secretory protein.
ACCESSION X68699
VERSION X68699.1 GI:53809
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 3777 to 5332)
AUTHORS Svendsen, P., Laursen, J., Krogh-Pedersen, H. and Hjorth, J.P.
TITLE Novel salivary gland specific binding elements located in the PSP proximal enhancer core
JOURNAL Nucleic Acids Res. 26 (11), 2761-2770 (1998)
MEDLINE 98256451
REFERENCE 2 (bases 7147 to 12653; 13952 to 17731)
AUTHORS Mikkelsen, T.R.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1992) T.R. Mikkelsen, Department of Molecular Biology, University of Aarhus, CF Mollers Alle 130, 8000 Aarhus, DENMARK
REFERENCE 3 (bases 7147 to 12653; 13952 to 17731)
AUTHORS Laursen J, Hjorth JP
TITLE A cassette for high-level expression in the mouse salivary glands.
JOURNAL Gene 1997 Oct 1;198(1-2):367-72
MEDLINE 9370303

FEATURES Location/Qualifiers

source 1..to 12653; 13952 to 17731
/organism="Mus musculus"
/strain="C3H/As"
/db_xref="taxon:10090"
/chromosome="2"
/map="Estimate: 69 cM from centromere"
/clone="Lambda YP1, Lambda YP3, Lambda YP7"
/clone_lib="Lambda-PHAGE (Lambda L47.1)"
/germline
/note="Allele: b"
misc_feature 3777-5332
/gene="PSP"
/function="salivary gland specific positive acting regulatory region"
enhancer 7147..8724
/evidence=experimental
exon 11778..11824
/gene="Psp"
/note="exon a"
/number=1
/evidence=experimental
exon 12626..14190
/gene="Psp"
/note="exon b fused with exons h and i"
misc_feature 12644-12652

Figure 5A:

```
/function=" consensus sequence for initiation in higher
eukaryotes "
misc_feature 13952-13965
/function=" M13mp18 polylinker"

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,
ACCESSION      M58708 L03370 L03371 L03372 L03373 L03374 L03375
VERSION        M58708.1 GI:145283
SOURCE         Escherichia coli DNA.
ORGANISM        Escherichia coli
                Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
                Escherichia.

REFERENCE      1 (bases 12653..13951)
AUTHORS        Dassa, J., Marck, C. and Boquet, P.L.
TITLE          The complete nucleotide sequence of the Escherichia coli gene appA
                reveals significant homology between pH 2.5 acid phosphatase
                and glucose-1-phosphatase
JOURNAL        J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE        90368616

FEATURES
Source          Location/Qualifiers
                12653..13951
                /organism="Escherichia coli"
                /db_xref="taxon:562"
sig_peptide     12653..12718
/gene="appA"
CDS12653       13951
                /gene="appA"
                /standard_name="acid phosphatase/phytase"
                /transl_table=11
                /product="periplasmic phosphoanhydride phosphohydrolase"
                /protein_id="AAA72086.1"
                /db_xref="GI:145285"

mat_peptide     12719 13948
                /gene="appA"
                /product="periplasmic phosphoanhydride phosphohydrolase"

mutation        replace(12659.. 12661, "gcg changed to gcc")
                /gene="appA"
                /standard_name="A3 mutant"
                /note="created by site directed mutagenesis"
                /citation=[3]
                /phenotype="silent mutation"
mutation        replace(13934..13936, " ccg changed to ccc")
                /gene="appA"
                /standard_name=" P428 mutant"
                /note="created by site directed mutagenesis"
                /citation=[3]
                /phenotype=" silent mutation "
mutation        replace(13937..13939, " gcg changed to gct")
                /gene="appA"
                /standard_name=" A429 mutant"
                /note="created by site directed mutagenesis"
```

Figure 5B:

```
/citation=[3]
/phenotype=" silent mutation "

DEFINITION pBluescript II KS(+) vector DNA,
ACCESSION X52327
VERSION X52327.1 GI:58061
KEYWORDS artificial sequence; cloning vector; expression vector; vector.
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 17732 to 20623)
AUTHORS Thomas,E.A.
TITLE Direct Submission
JOURNAL Submitted (20-FEB-1990) Thomas E.A., Stratagene Cloning
Systems, 11099 North Torney Pines Rd., La Jolla, CA 92037, USA
REFERENCE 2 (bases 17732 to 20623)
AUTHORS Short,J.M., Fernandez,J.M., Sorge,J.A. and Huse,W.D.
TITLE Lambda ZAP: a bacteriophage lambda expression vector with in
vivo excision properties
JOURNAL Nucleic Acids Res. 16 (15), 7583-7600 (1988)
MEDLINE 88319944
REFERENCE 3 (bases 17732 to 20623)
AUTHORS Alting-Mees,M.A. and Short,J.M.
TITLE pBluescript II: gene mapping vectors
JOURNAL Nucleic Acids Res. 17 (22), 9494 (1989)
MEDLINE 90067967
FEATURES Location/Qualifiers
Source 17732 to 20623
/organism="synthetic construct"
/db_xref="taxon:32630"
CDS complement (18967..19827)
/gene="Amp"
/product="b-lactamase"

BASE COUNT 5449 a 4847 c 4902 g 5424 t
ORIGIN
1 TCGAGAGTAT CTTTGTCACTG TGTGCCTCCA ACAAAAGGGGT ACTGTTGCC ACATAGAAAG
61 ATCTAAACTA ATTAATTAAT CCCTCACCCG CAAATCTTC AGTCACTAAG TTAGCACGAT
121 TGTGAACAA GTTCTCCAAA GGAGAGATAC AGATGAGTGC GTATAGGGTG GACCTGGCTG
181 CTGAGGAGAC ACCTGCATCT GACTAAGAAG AGCCACGGTG TTAGTTGAAT GGTGTGGAGT
241 AGGGTGGTTC TGTGGGACAG TAGAAAATCG AGAGGCATGT GCCGTTAGT GAACTGATGG
301 AAGCTACCCC AAACCGACAGA GATTGTCAGT CAGGCCAATC CGTTTCGAGT TTGATGGGCA
361 GCCGGACAGT GAGACAGACA CACCTACTCA GTTGGAGGAA GGATGAGAAC AATGGCCAGC
421 AGGGATTGAG AGACCCCTGAC AGGCGCAAGG CCCTAACACA CACACCTACC ACCTCACCTG
481 ACAAAAGCTGC CAAAGACCAA AGACTTGTTC TCCATTAGAA ATGACAGCTG GCTTGACCCG
541 ACAGCATAAT AAGCAGAGTG TACTCTGATT GGAGAACTTT AATGTGTTTC ATTCACTGATT
601 ATAAAAGGAC AGTATTACAG ATTTTGTGT ACACTGCTGT TACATGTGGG GCAGTGTGTC
661 TTTAAGTAGG GTAAAGTACT CTTTAAAAAT GGGTCCTAGA TATTTTTTCC TTTAACTCAA
721 GTCTCTTACT GTTTAAATGA TTTTTATTTT GTTTAATATG GAGGAAAAAG AAGCGTAAAT
781 GGACAATATA TATTTAGAGA AAGATGGTTA GCTGTCAGAA AAATATGCAA ATCAAAATCA
841 CACCAAGACT GCAGCACACC CCTGTCAGAT GGCTGTGATC AAGAAAATAA ATGACAATGA
901 GTGGTGGTGA AGATGTACTA AAGGGAAACA CACACACACA CACACACACA CACACACACA
961 CACACTGGAG CAACCACTGT GGAAATCAGT ATGAATGGTC CTCAAAAACC TGAAGATAGA
1021 GCGGGCGTG GTGGCATAACA CTTTTATTCC CAGCACTGGG GAGGCAGAGG CAGGTGGATC
1081 TCTGAGTTCC AGGCCAGCCT GGTCTATAGC ACAGGTTCTA GGACAGCCAG GGCTACACAG
1141 AAAAACCTG CCTGATTAA ACCAAACCAA ACCAAACCAA ACCAAACCAA ACCAAACCAA
1201 ACCAAACCAA ACCAAACCAAG ACCAAACCAA AACACTGAAG ATAGAACTTC AGTATCCAT
1261 TCCTAGATAT ATACCCAATG GAGACTAAGT CAGCAAGACA CCTGCACAGC CATGTTCACT
1321 ACTACACTGT TCACCCACAGC CAGGCTGTGG AACCGCCCTG AGTGTCCATG ATAAATGAAAT
1381 GGATAGGTAA CTTTCAAGGT AAATGGACTC TGCTGTGTAC ATGCCTCACA TTCTGTTTAT
1441 TCATTTCTT TTATGAGGTG TCCATTCAAG AGTCACATGG TAGTTCTATT TTCAGTCTTC
1501 TGAAGATACT ACACTGGTCC CCACAGTTA CACTTTTATC AGCAGTGAAT AAGGGTTCCT
1561 CTATCCTTAC CATCATTTGT TGTAATTTT CTTGATGACC CTCTTCTGA CAGGGATAGG
1621 ATGTAATATC AGTGTGAGGA AGTACAACCTT GTTTCTAAG TATTTATTGG CCCCTTGCAT
1681 TTCTTCTTT GAAAATGTC GGTTCCGTGAC ATCTGCTCAG GTATTCAATTG GATGTTGTT
```

Figure 5C:

1741 CTTTGGTGTGTT TGAGTTCTTA TGAATTCTAG ATGTTAAATC CCTGCCTGTG GTTCTCTCCC
1801 ATTCTGTAAGG CTGCCCTCCTC ACCCTGGCAA TTGTTGTCCT TGTTTTGCAG AAACCTTTGA
1861 CTTCATGGAA TCTCATTGTT CAGTTTCCC TCCTCTGCTA TAGCCTGAGC TAATGCACTG
1921 GTTTTACAG AGCCCTGGTC TATGCCCTTA TCCTCCCTG GCAGCTTCGG AGTTTCATTT
1981 CTTACATTAA GATCTTGAT CCACCTTGAA CAAGTTTGAG ACAGGGTGA GAGATACGAA
2041 TCTAGTTCCA TTCTTCCATA TGTGATCCTA GTTACATAG CATCGTTGGT TGAAGAGGTT
2101 TTATTTTATT TTTAAATAAT GTGTCACTAA AAACGAGGTG GTGTTAGCAG TGTGGATTG
2161 TTTCTTGTG CTTTGATCTA CAGGTCTTGT TTTGTCTCAG TCTCATGATG TTTTATGCT
2221 ATGGCTCTGT CATACTAGT CAGGTCTAGGT ATTGTGATAT ACCTCAGTA TTGCTCCCTC
2281 AGACTCAGGT TTGCTTGGC CAGGAGTCAT TTACTCAGT GCTCTTAGAG CTCCCCCAGC
2341 ATGTTAGCTGC TACTATTCTT AGTTGATAAA TCAGGAAACT GGGGCTCAGA GAGATTAAC
2401 GTCTTGAAC ACTTCTGGGG AGGTGAAACG TGGAGACACT AAACGTGTT TACCCGTAC
2461 TGCTCCAGTA GCTGTCGGGT GCTGGGCTAC AGCAAAGCAC CTATACTATA TATTACTCAG
2521 GAGGTGGAAA AACTCAGCCT CCCTTGGGGT TCCCAAGCTC CCAGGTGTCC AGTCACTGCT
2581 GGAAACCTCA TGGAGTCTGA AAGGAAGGGT TGAGGGTACA TGGGGCAGCG ATGAGGAGCC
2641 TGGGGCTGGG ATCTCCCAAAC CACCTGGATA TCCAGATGCC ACTGGGTCA GGGGAGTTGG
2701 GAACAGAGTT GGGATGTCCA TGGACCTGTG ACAAGGCCAG GGCCAGGGGG AGGATAACTC
2761 TGGCTTACT AATTTGCGAA AGTCCTTAGC TTAGCAGCAG TTGTCCTGGGA GCACAGAGGG
2821 GCCTTCTGTA AGAGGCTCAG CGAGTGCCTGC TCTGTTAGCG AAGGTCTTCT CCATGTTCCC
2881 CATGGTGGTT CTTGATGAAA GAGACAGTCC TTGGCTCCAA ACTGGTTTAT TGATTGTTCA
2941 TTGTGGAAA TGGGTGCACA CCACCTTCTC AGGGTGGACC AGAGATCAAA TACCTTTGC
3001 AGGGAGGAAT ATCTGGGAAG GGACGCTTAC TGGCTAAACC CTCAGGGCCT CTAGATACAT
3061 CATTAGCATG GAGAACTCTG TTCTGGCTA CATGACCACAA GGCCACATT CCACAAGCCA
3121 CATGTGGAA GTGTGGCACA TGTCTAGGC CAGGAATCTG GTAGGGAGCG TGGAGCCACC
3181 TACCATCCCA GGTGGGTGCC TGGGTGCCAG GGACCTGAA CCCGCTCAAC CTTACCAAGT
3241 TTCCTGGCAG GGTCCACTGT CCTACACAGA AGCTGGAGGA GGTGTGAGGG TTGTGTCTT
3301 GTGGAATGTC CCATGCTGCT TGGGGCTCAG TTTCTCCACC TGACCTCAT TGGTTGGGT
3361 ATAAAAAGTG GGGATACTTT ATTATTCTCT GACTCGGTCC TGAGGAAAAAA GCATCGTGGC
3421 AGTCCAGGAA CCACACCCCTG AGGTCTCTGC ACTGAAGGGG CTCCCTAAAGT CTCTGGAGTC
3481 TCTCCCTCTC ACAGAGCTGC CAAAGTCTAG GTTCTTTGA GGATAACAGA GCCATGCTTG
3541 GTAAAGCAGAC AACAGCATTT GTTACTCAAC CTTCTTTTG TCAGCTCCCT CTTCATAAAC
3601 AAGTTGAGAC ACCATGCTGG CTTGAGGAAG ACTTCTAAAG CCAGACAAC GTGCAAGGAA
3661 GAAGAAGAAG GGGCAAGTGG AGTTAGCCTG GATGTAGCCC TCAAAGTCTC CAGAGACCAG
3721 CCATGAAGGC TCAAGTGGAG GGCAAGACCT GCAGCAGCCA AGCATCTGGC AGGAGAGGAT
3781 CCTGGGAACC CCTCTACCAT GACACACATT CTTCTGCAG GTCACACTTA ATAGGCCATT
3841 TCTTATTGAG ATCTATCATG GTGTTCTGTG CGAGATTAAT GAGGTGTTAT GCTGCAACA
3901 GAAAGTTATA TAAAAACAAG TCCCCCCCCC TTGTCACTGC TGCTAAGAAT GTAGCAGAAA
3961 TTGTCTCAAG TGTCTCTCTA ATCAGAAACA ATAAAGGTCT CTTGGATTG AAGCCCTCCA
4021 GTTTCTCTCT CTCCTGCTGA GCCTTGGACA CCCATACAAA CCTCCTGGAT GCTACAGCTC
4081 TGGGAGAGA CTCCAAGGTG GGGAGAGACT GATGGTACAA AAGCAAATAA CTTGTTGGG
4141 GGTACACCCA CCTCTCTGCC TGTGTTGTTG CTGAGTCAG TCTGCAAGAC AGGCCCTCAG
4201 TGGGTCTTCC ATGGGCAACA CGCAGAGGGG GGCAATGGAT GGAATACCC ACACCCCTGGT
4261 TAGTTTACCC CGGCCATGCT CTCTGCTCTT CATCTCTCTG CTGCCCTCTG CCACGGCTTT
4321 CTCTGCAGGA ATCATATCTT CATATTGGCC CACAGGTGTT CTCTCACCC TAGCTATGAT
4381 GTTTACTTTA GAGTGAACCTT AGCAGGGCTG GTGGGAATGA GTTCTAGAAG GCTCACGGAG
4441 ATGCTAGGGA AGAAACGTCT TCTAACTACT GAGGTTACTA AGTTCTGGT GGTTGTCTCT
4501 GCCTTCCCT TGTTAAAGTC ACCTTGAAGT TAGTGCAGAA GAAATCAGAG CCCAGTCACA
4561 GAGTAAATAT GTTCTGAAAG ATTTCTTTG AGTGCCTCAGA ATCCATGACA TTTCAAGAGC
4621 CCTCTTGTG CTTTAAGTCA TTGGGGTTG TATCTCTGC TTGATGTATG TGTGTGTGTT
4681 TATCAAAGAG TGAGATGGTT ACATAAGAGG TGCTCTAAAG GACAGAGAGG ATTTGCAATT
4741 GTGGCATGTG ACATCTCTAG GCCTTGTCTC GGTGCGAGGA GGAACGTGATG CAGAAAAGAG
4801 TAAGAGGTCA TTCTCTGGAG GCTGTCACTA TAGAGGAGAT CTTACAGTGC ATTCCCTCCT
4861 CCAGGCCCTG CCTGAGGATA GACATGTGCT GACTGCAACT GAAACAGAGG CTTGGGATGG
4921 AGAGTTAGGT TCACAGAAGG GAGGGTGGGA GATGGATGCT TGCTGGGTT TGGGTCTCAT
4981 CACCAGCTCC TGACCACCCG GTCAAGCCCAT GTGCTTATTG CATAGCTTT TTTGCTATG
5041 TTTACTCAGT GTGGTGTGTT TTGGGACCCA GCAGAAGCCA GTCCCAGGCT GACAGCTGTG
5101 GATACACAGG GCAGCATGAG GGTCTCTCAGC CTGAAGCAGT CAGGCTGGCA GAAGAGAAAG
5161 ACCAGCACAC ATTCCCTCAA CCAACTATGT CTTGAAAAAC AAACATATTA TATCACATAT
5221 ATTGCATTAA TGAGACAGCT AAAATGTACT CGGGTAGCAT GACTCCAGGT GGGGATATCT
5281 GCAAGTGCCA TGAGTGGCAG AGGGACAGCC AATGTGAGGC AAGAAGGAAT TCTGGCTCAA
5341 CACAGCTTAG CTCCTGGTG TTGGTTCAA CTTTGAGAGT TTGACCCACAA GCACTTATT
5401 TTGACATAT TAAACAGAG CACAACTTTG GGAAAAAGTT TTCTTATGAA AATTATCACA
5461 ATAAAGCTTA AGGCATGACT ACATTAACAT GCCTTGCAGA AGTATATGTG CCCTCTTCCA
5521 CAAGAATGGT TCTATTGACT GAGAAATAAT GTTCAGGATA AAGATCCAGG AAGAAAAGAT
5581 CAGGGATAAG TAAATACAA AACTCTTTG CAAAGTACAT AGACCCCTTT TCATAACAAT

Figure 5D:

5641 GGGTTCTATT GACTGACAAG CACTGCTCAG GAGTTGGAA AGAGTCTAGC ATAAGCACGA
5701 TAGCCTGGAG ACTCTAGTGA GGTCTAGTCT TACAGACAGC AAAAATCACC AGGTTACAAA
5761 CTACATTCACT TTCCAGTTT CTGATCAGGC ACAGGTATGA ATCCCTCTG TTGAAGAGAA
5821 AAGTCCATGT GTTTAAAATA TCTGGTTCT CCAGTCTAT TAGCGAGAAG ACTTGAGCCC
5881 TATACAACTC CCACCTGGAG TGACATCCTG TCTTCATGGT ATATTACATA CCTAGACACG
5941 CTCATCTCAC AGACTTAGGA CTTTGTCTC TGATCTCCAT TTCTGATCCC ACTTCCACCT
6001 TTGAGTGTGAT AGTGTCTATT TCTTCACTGC CTTGGTGACA ACCATGTTAT CCTCTGTGTA
6061 TTTGAGTGTGAT ACCATTTCA GATTTTACCT GTATGCAAGA TCACAGTC TTTGTCTTC
6121 TGTCTGGATG CATGCTAAC TCTACACAA AACCCCTCCC CGTCACTCAG ATCTTCCTCC
6181 ATTAACACAT ACATGGTGCT GAAGAGGCTA GGGAGCTTC CTTCACTGGG GAGCTAGCTG
6241 GCTATTGGGC CTTTTGACT GTCCAGGAAG GCCCCCAATT GCTGAGACAA GAACCTAGAT
6301 TCTTCATTAT TGACTCTAAC TCATGTATCA AGCAGAAGCT AATGAATAGT TATCAACAGG
6361 ATCAGAGGTT CCAGTGTAAAG ACACCTTGAC ATGAAAGAAC GGAGGAAGGA CAGATGGATG
6421 CATAAAAGCA GGACCACTGC CCCAGGAAGG TCCTGAAAC TGATGCAGGG CAAAGGACAG
6481 GTTATAAACCC AAATCTTAGG GAGTCAGGAA GAGCACAGAG GAGCTCAACC AACTGACCAC
6541 TGCTTAGGGG CTACCAACCC AATCCTCCCT GTGGGAACAG CTAAGCTATC AGCCAAGGGT
6601 AATAAACAGG CAGGACCTGT GGATGACATG GAGAGCATAG GGACCCCTGGG TCCAGCTTT
6661 AGCACCTGCA CTCTCAGGAT ACTCCACCAT TGTGCTTAGG AGAGCCTAGG GATACTGGGT
6721 CCAGCCTTGC GTACCTTCAC TCTCAGGGTA CCCCCTCACT GTGCTTGGG GAGCCTAGGC
6781 ACCCTGGGTG CAGCCTCAG TACCTGCCTG CTCAGGACAC CCCACCAATTG TCTCTTGGCC
6841 CGTCTCTCTC TCCTCTTCCCT CCCTTCTATT GTCTCTTCTC TTGTTCTTTC TTGACTCTCC
6901 TTTCCTCTCA CACCCCTCACT CTAGTTCTCC CTTTCCCTCT CTGCATCACC CTATTCTCTC
6961 TGTGGTCCCT CCACTTCTC TTATCTCTCA TGCTTCTCTC CTTCCCTCAA TACTTGTAC
7021 CCACTATACT TCAGGGGCCA GCTCTAGTGA CAAAGCTTT AATAGCAAGA CTCTCAGATC
7081 TCCAACGGCT CAGAGGAGCC AGACCCACCA AGAACTCTCT CCAGGTCCAA TTTCAGGTT
7141 CTTCGAAAGC TTTCAGCAAA TGCTCAGGGA ACATGCCACT ACAAGAAGA TGCAAATTCC
7201 AGTTGAGAGT GGGAAAGGCC CTTGCGTAGG TCCCCATCTTC CAGGCCAAGG TCAGAGGGC
7261 TCTGTATAAT CCGGATTGAC AGGGCTCAGA ACAATGTTT GTTTTAAGG TTTATTATT
7321 TTAGGTGTTA GTGCTTTCAC TTGCTGACCA TTATGTGCT CATGTTGTTG CAGGTTCTG
7381 ATGACAGTAG AGGAGGGCTT TGAATCCCTG GGGATAGGAA GTACAGGAA ATTATAAGCT
7441 GCTTTGTGGG TCTCTAGCT TTCCCAACAG AAGTGAATGC TCTTCACAC TGAGCCATCT
7501 CTCTAGGCC AAGAGACATT GCTTTATGGA TATAATTGTT TGTTGTTGTC AACATTGAGG
7561 AAAGGGAAAT AAAAAAAAGG CTTCAGCCGC TAAGGTGTA CAGTTTCACT AATTGCTACT
7621 TTTAGTTGTT ATAAAATGGC AGGTGCTCA ACATTATAT ATACAAAAAC TTCCCTGCTG
7681 GTGGTTCAAC TGTGAGAACT GGGGTAAGTG GGTGAGTTCT CTTTTCTGT CTCTGTCTCT
7741 GTCTCTCTCC TTCCATTCTT TCTTAAAGGA AATAAACATT GCAGCTGGGT TATAGCTCAT
7801 CAATATGGAA GTTACAGAAG TGAAAAAAGG CATTGCCCTTG GTGGGTGGTG TTACCAAGCTG
7861 ATTTTGGTT GTCTGCAAG GAGGTCTGGG GACTGGCTGC TCTGCTCTG TCTGTATGAG
7921 TGAGGGAAAGT CTGGGGAGCA GATTCTCTAA CCTTCAGCC CCCCCTGGTTC CTGAGTGAAC
7981 CCAGCCTCTC TTGTCCTAGT AGCTTTTCC AAAACAGGAAT CTGAGTGGTG ACAGGGAAACA
8041 AGTACCAAGGC CATTGCTTAA GTGCCAGGGT TAGTGAGGGC AGGAAGCTGC CATAGCTGGG
8101 ATTAGTAGTT GTATTGGATG TAGGAAGTCC TATCCTGGG CAGCTAATCC TTAATGCTTC
8161 ACTGGAGATT TTCAATGAGA AATTATGCC ACGGCCCAT A TGCCCCCATC CTTTTGTCTC
8221 CAACAGCCAA GTATTGCCA TTAGAGGAGA CTTCCTGTAC ACTTGATGGA TGCTCATTCC
8281 AAGGTGACTT GGGGCAGTCA GTACAGACTT GGGATGACCT CTGACAGCCT AACCTCTCCC
8341 CAACAAGGGC CCTCTATGTT TGCTATGTA TGAAATGTCA GACATTGTCA GGAGTGTCCG
8401 CAGCACAGCC TGCCCACTGT GAGGGCTCTC ATAGGTTTCC CACTGTCTTA TCTACACAGG
8461 GATAACGAGG AGGTAAGCTG CAGTCTCCAG TCTCACTTCA CAGAGGAAGA GATAACCCCA
8521 TCCCAGGTCAT TGAGGCCAGC AGTGGAAAGA ATGAGGATT GAACCTCAGGT CTTCCAAGTC
8581 CCATTGATAG CATCTCTCA CAAGTCCCTT GCCACCCCTCA CGATGCCCTA GACACTTGCC
8641 TGCCCTTTAT ACTAAGGAGA TGCAAGGTACA AGGGGTTAC CCAATGTAGCA GCTGAGGCAG
8701 CTGGGGATAG ATACCAGCAG CAGGCCCTGAT GTCACCACTC TAACCTCCAGC ATCCCCAGTC
8761 TGTGTTCTG GAGTGTGAAA ATCCCTACTT AACAAAGATTG TGCAACAGTC CTTGGCTCTG
8821 TGACCCATAG CTGGAAACAG GATTCTCATT GATTTGTGGA ACATGGTGGC AGCCAGCCAA
8881 AAAGAGGGTC TGCATACAGA AGACACGTGT GGCAAGGCCA CAGCAGACTC TGACTACCTT
8941 AGCTTACAGA ATTACAAGGT CATAATGTCC TCTGCTTGG TCACCTCATG TTAAGGACAG
9001 GCCCTAATGA AGATGGGCCA GAAGACTGAA GGAATGGCCA ACCAATAACT GGCCCAACTT
9061 GAGACCCATC CTACAGGCAA GCATCAATT CTGACACTC TAATGATACT CTGTTATGCT
9121 TGCAGACAGA AGCCTAGCAT AACTATCCTC CGAGAGGTCC ACCCAGCAAC TGACTGAAAC
9181 AGAAAAAGAT ATCCACAGGC AAACAGTGGA TGGAGGTCAAG GGACTATTAT GGGAGAGCTG
9241 TGGGAAGGAT TAAAAACCCCT GAAGGGGATA GGAACCCAC AGGAAGACCA ACAGAGTC
9301 CTAAGAGACC TGTGGGAGCT CTCAGAGACT GAGCCACCA CCAAAGAGCA TACACAGGCC
9361 GGTCCGAGGC ACCTGGCACG TGTGAAGCAAG ACATGCAGCT CAGTCTCCAT GTAGGTCTC
9421 CAATAAGCGG TAGCCTGACT GCAGTATCCA ATCCCCAACAA GGGCTGCATA GTCTGGCCTC
9481 AGTGGGGGAG GATGCCCTA ATCCTGCAGA GACTTGATGA GTGGAGAGCT ATCCAGGGGG

Figure 5E:

9541 AACCCACCCCT CTCTGAGAAG GGAATGGGGA TGGGGGAGGG ACTCTGTGAA GAGGGGACAA
9601 GGACAAACAA GAACCTCAA TAGGTCAGGC CCTAAAGGCT TGCTAAGTAG CAGTGGCCA
9661 GCTCTGTCCT GTTCCTCAGC CCAAGGCTCA GCTCCACCT GTTCTGTGT TTTCTGGCT
9721 TTTCATGGC CTAGGACTTG GTGACCAGTT CAAACAATGG GGCGTGTGGA AGACACAATA
9781 TACAAGACTA GGGACATTCC TGTTCTGCTG ACTATCCATA GCCTGTGTA GGTGGAAGGA
9841 CCCAATCACT GGATTCTAC CCTTGCACAA CCTTGACAGC TGAGGGCCTC TCAGAAACCT
9901 ATTTCTTCA CTGAAAAATG AGACTCTCAA ATGAACTGCG TGACAATCAT CAGGCTTATT
9961 AAAGAGGTGT ATCTAACCTG ATGGCAAGC AGACAGCAGG CAAATGTCTG TATCAACCTC
10021 TAGGAAGGAC AACAACTGCT CACTGCTGCC CCCCAGGAGG CCATTGCTG AAACAGCTGC
10081 TCTCCTGCTG GTGCACAGGC CCTGCCTCT CATTGAGCC ACAGCCCCCTT CCTGTCTGAA
10141 CCTCCTGTC GGTCACTGGG AAACAGATCA AGATGGAACA GGACAGCTCC TGATGGTAA
10201 TAAAAAACAG TGGTCATGGC TATTCA TAGG CGTTTATGCT TCTTCAGTCC ACACTGTGAA
10261 GAGCTGTGGG CATGAACCAC AGTGTTCGAG GTAGAGTTGG GTTCTGAAA TTCACAGTGG
10321 GGTGAGCTCA GTAAATGTGA GCTGGAGGTC ACTCGTGAGA CACACAGTCC TGCTGCTTCT
10381 GTTCCCAATA TCCTGAGGAG ACGACACATC TACTTTGTT AGAGGCCACA GTCTAGTTGA
10441 CCTGAGAGTT ACCAGTTCT TATTGTGTG TGTGTGTGT TGTGTGTGTG TGTGTGTGT
10501 TGTTGTTCGT GTGTGAGTGC AGGTGCACAT ATGATAGCGT ACACGTTGAG GTCAGAGGAT
10561 AACTATCAG CGTTGCCCCC CCTACTTTT CCTCGGACTC TGAGAACAA ACATGGGTCC
10621 TTATTCCAGG GGAGCAAGTC GCTGTTGGCT GACACATCTT GTCACATAC ATTTTACCTA
10681 GACAATGGAG CCTCCATCAG AGTATTACTT TAGCTCTCA CCGATGGCAA TGCACCACCT
10741 CTCTACCCAC ATAGGAGTTG GGTCTCCACA CACCCCCACA CCCCTTCAC CAAAACGTT
10801 TCAGTTACTT TATCTGGTAA AGTTCATCAG AGAATGAAGC CAGTATTAAAG AACATGGAAT
10861 CATTGGGAA CCTGGATCTA GCAATACCCC ACCCTAGATG GAGTTGCTGA GTTTTACCT
10921 CAGATTATAA TTCCCCCCTA GCTTCTATGG TTTATTCTGA AACCAGGGGA ACTCGATTCC
10981 TCCCTTGGG CAACAGACAT CCTGGCTTGT GAATTCACAT GTCATCTACT GCTAATCCAT
11041 TGTTAGTATG TGGCTCACAG AGACACACTA CAGTCATGGC CAATGTCAAG GTAGGACAGA
11101 TGTGAATCAT TCCCCCAGTC CTGCTGTTT CATGACTAAC CCTCCTCAGC ACAGTGACCA
11161 TGAACCTACT TTTCCCTCTT TTTTATTTT AGAATTGCTG GAATTCTTA TTTTGAGAAA
11221 TAATAGCCCTT GGGCAGCATT AAACAAAATC ATCTAGAAAG CTGGTTAAA ATACAGATGG
11281 TTGAGTCAGT GAAAGAGTGA GGAATGTCAT TATTGGCCCC TCACAGAGGC TGGCTCACTC
11341 CAGCAGAGGT GTTGAAGCT CTTGGACACG GGTCAAGGTGCA ATAGGAAAGG TNGTCTGGG
11401 CACTGAGAAC CACAATTGAA CAAACAGAAC TGTTGCTTT TTTTTTTA AATGAGTTCT
11461 CAAAAAAATGA CTGGCTAGCT TAGGCAAATA CTTCGAGCCA ACCAACAGAA ACATTCTTCC
11521 ATTGATTCA TCTGGATCTT CTTCTAGAC AATACTGAAC TGACCCCTTG TTGGCAGTCT
11581 CAAGTTTGAC AACATAGGGC TTGAACCTG GCACAAGGTC CATCACTGTC ACCCAAGCAT
11641 CCTGGGTGAC CTTTGGGTTG GAATATCTTG GCTAACCTTA GATATTCTT TTGGAGTATC
11701 TTTAGAACAT CCAGGAAATA GGGCTTGATT CTCATCTTG GACCACAATA TAAGTCACCC
11761 TAGAATCCA GGAGATCGTG CAGAGAAACA AGGATCTCTC TCGTGTGCAT CCTCTCTCAA
11821 AGCAGTGAAGT AGTGAACCTCA CTAAACTGAG TTCCCACATG AGATGCCACA GGAGGCTTTG
11881 GGGCAAGAAC CAGAGGGAAAG CCACTGTTG TGTTGACTC TAACAAATT
11941 AAAGACATAG ATGACATTGT GTCAGACTAA CAACAAACCTA GACTCATGTG GGTCTGTTT
12001 AGGGATCAGA TTTTATTCAT CAATGACTTG TCTTAGTGTAGAGAAAGG CTTCTACTG
12061 GAGTGTAGGC TCAATAATGA CAGAAGAGAT AGCTATTCTC CCTAGGGACT GTGCTGCTCC
12121 AAGTTGGTG GAGAAAGGCA GTGGGGAACC TAGATGTGCT CTCTGGGGAG GGGGTCTGAA
12181 GCTGGCTTC TAGAAGGTGT GAAGTTTGC TGAAACATCT AAACAGAATT ATAGCTTAGG
12241 AAAGTGAGCA GGCAAGGCAG GGAATGTTGTCATGATGTA ATGTACATGA ATATATTATG
12301 TTATAGATAC ACACACATTG GAACCTCATT TGCAAGATGAC AGAAAATAGG TTATTTGCC
12361 TCTCTTAAC TCTAACGACA ATGACTTCCA GTTCCATCCA TTCTCTGAAA TGCCACAAATT
12421 TCATTTTCA TTGTGGCTGA ATAAAATTCC ATTGCAAGACT GGGCCCTACT TCATCCACTC
12481 CTGAGGGCAG GCATATCCCC TGGCTCCATT TCTTACCTAT TGTAAGAGAGA AGTGCACACTG
12541 TCTTGTGAA AGGCAAGCCT GAGAGAGGCA GGCACAAATT GTGGGTTTTT GTTTCTCTT
12601 CCTGCTATGA CTCTCCATTG TGCAAGACCA AGATGCGATA AAAGCCGCCA CCATGAAAGC
12661 CATCTTAATC CCATTCTTAT CTCTTCTGAT TCCGTTAAC CCGCAATCTG CATTGCTCA
12721 GAGTGAAGC GAGCTGAAGC TGAAAAGTGT GGTGATTGTC AGTCGTATG GTGTGCGTGC
12781 TCCAACCAAG GCCACGCAAC TGATGCGAGGA TGTCACCCCA GACGCATGGC CAACCTGGCC
12841 GGTAAAATG GGTTGGCTGA CACCGCGCGG TGGTGAGCTA ATCGCCTATC TCGGACATTA
12901 CCAACGCCAG CGTCTGGTAG CCGACGGATT GCTGGCGAAA AAGGGCTGCC CGCAGTCTGG
12961 TCAGGTCGCG ATTATTGCTG ATGTCGACGA GCGTACCGT AAAACAGGCG AAGCCTTCGC
13021 CGCCGGGCTG GCACCTGACT GTGCAATAAC CGTACATACC CAGGCAGATA CGTCCAGTCC
13081 CGATCCGTTA TTTAATCCTC TAAAAAACTGG CGTTGCCAA CTGGATAACG CGAACGTGAC
13141 TGACGCGATC CTCAGCAGGG CAGGAGGGTC AATTGCTGAC TTACCGGGC ATCGGCAAAAC
13201 GGCGTTTCGC GAACTGGAAGGGTGAAGCTTAA TTTTCCGCAA TCAAACCTGTG GCCTTAAACG
13261 TGAGAAACAG GACGAAAGCT GTTCAATTAC GCAGGCTTAA CCATCGGAAC TCAAGGTGAG
13321 CGCCGACAAT GTCTCATTAA CCGGTGCGGGT AAGCCTCGCA TCAATGCTGA CGGAGATATT
13381 TCTCCTGCAA CAAGCACAGG GAATGCCGGGA GCCGGGGTGG GGAAGGATCA CCGATTACAA

Figure 5F:

13441 CCAGTGGAAC ACCTTGCTAA GTTGCAATA CGCGCAATT TATTTGCTAC AACGCACGCC
13501 AGAGGTTGCC CGCAGCCGCG CCACCCCGTT ATTAGATTG ATCAAGACAG CGTTGACGCC
13561 CCATCCACCG CAAAAACAGG CGTATGGTGT GACATTACCC ACTTCAGTGC TGTTTATCGC
13621 CGGACACGAT ACTAATCTGG CAAATCTCGG CGGCGCACTG GAGCTCAACT GGACGCTTCC
13681 CGGTCAGGCC CATAACACGC CGCCAGGTGG TGAACCTGGT TTGAAACGCT GGCCTCGGCT
13741 AAGCGATAAC AGCCAGTGGA TTCAAGGTTT GCTGGCTTC CAGACTTTAC AGCAGATGCG
13801 TGATAAAACG CCGCTGTCA TAAATACGCC GCCCAGAGAG GTGAAACTGA CCCTGGCAGG
13861 ATGTGAAGAG CGAAATGCGC AGGGCATGTG TTCTGGCA GTTTTACGC AAATCGTGA
13921 TGAAGCACGC ATACCCGCTT GCAGTTTGTG AGGTACCCGG GGATCACAAAC TTGCCCTCTG
13981 AAGAGGAAGA ACAGAAGGAT GCCACAACTC TCCTGCTGG TACTCTCCAG TGTTTCATC
14041 TTACTTCTGA TGGCATTTCC CTCTAGAAAG TGCTACTATC ATCCACACAT TTCTACCTGA
14101 GACCACCCAA AGGACCCCTCC CAAATTCTCT TCCTCTCTGA GTAGTCTCCA CACCTGTTAC
14161 CACCATCCCA GAATTAAAAT CCTAAGTGCCTA CTCTGGCGT TGACTTGCCT CAGTCCTTGC
14221 AATAAGAGTT GTTGGCAGTG CCAGGCGTGG TGGCGCACGC CTTAATTCC AGCACTGGG
14281 AGGCAGAGGC AGGCGGATTT CTGAGTTCGA GGCCAGCCTG GCTCTACAGAG TGAGTCCAG
14341 GACAGCCAGG GCTATACAGA GAAACCCCTGT GTCGAAAAC CAAAAAAAGTTTT
14401 GTTGGCAGAG TGTGGTTAT ATACCAGGTG GAGATTCTAA ATGAGTGGCT GAAGCTGTAG
14461 CCAGAAGGAA CTTAGAGGAT AGCTCATAAC TTAAAAAGAA ATGTAGAGAG TAGCAGAAAC
14521 ATTGAGAGAG TGGGCACACA GCCACTGTGT GAATGTGGCA GACACAAATC CAGCCAGCTA
14581 TACATGCATA AGTGTATATT GGCGCCATCC TGACTGATGA GACACAGGAA AACAGATAGA
14641 CGGGGTTAGG TGGCCATGGC CTTCTCTGCC TGCCCTTCC TAAGGGTCAT CTCAGACCT
14701 TATGCTCTCT TAACTCTTCC ATTGCTACTT AGCTTCTAGA TATCACCTCC AGATTAGTCT
14761 CCTTGGGTAC ATCAGTGATC CTGGTGTAT CCAGGGCTTC CTGATTCCAT CTTTGTCTA
14821 GAGGCTGCAA CTAAAGAGGT CTTCTTAATA CTTCACACCC TGATGCCAAA AGGAAGACAC
14881 AGAAGTTCAC AGAGGTGAAG TGATTCTATG AGGACATACA GTGAGCAAGC ATCAGGGTCC
14941 GGATTATCTG ACTCTACTCT AACTTTTATG TAAATGTGCT TTATGCCATT AACACTGTCA
15001 TCCATGTGCT TCAGCTCTGG GAGACTCCCA AGCACTCTTA GGACAAAGCC ACAATTAAAGG
15061 GACTCTGACA CTCTGCATTG ATTAATTAGC ATGGGGTCT CTATGTTTCC AGATTCTATGA
15121 TTGTTTCACT TTCCATATAG GCTATGAAGG GTGTGAGGAA ATTTTTGGG GACAGAATTG
15181 GAGGCAATTCC ACCTCTCTCA GGAAGCCTCT ATCTGGAAAA GCTTACAAC CAGGGACAGT
15241 AACTGTAGGC CCAGTCCTTG GTGTCCAAAAA TGGGTTTTAT GGGTTGAATC TGCAAAGCCT
15301 TCCATGTGCT CAAAGGTTT AACTGGAGC CTCCTCTGG TAACACTGTA TTGGAGGCTT
15361 TTGAGACTGG ATGCTTTTG GTCCCATGT TTGCTACATC ATCTGTCAAG ATATGACCCA
15421 GGCATGCTAC CAGCTACAC AGACTATGCC TCTCCAGCTT TCATGTTCTC CCCACCATGA
15481 TAGACTTGTAA TCTCCTAAAAA ATGGAATCAA AGCAAACCTT TCCCTGCATTA AGTTTTTTT
15541 TTTCTGTTAA GTGTTGGTC ACAGGGACAA GAAAACACTC AATACAGATA ATTAGTACCA
15601 GAGTTGAGGT TCATTGCTCT CGAACAGTTGG ATCAAATTTC TAGGGCTTTG GAACTGATTT
15661 ATAAGAGACA TGAGAAGAG TCTGAAGCTG TGGCTACAG AAGTGTCAAC AGTTTTAAG
15721 ATAGTTTAA TACACCATGG GAATTGTGAA AATCAGAATG CTACACACAA GGCAGACAGG
15781 AAAACGTGAG CATGGGGCGT GTGAGAGGGC ATAAGAAGGA ACTTAGGGGG AAATGAGCTA
15841 GAAGCCATTG GGCTACGTTA GGGAACGTGT GTGGCTGTG TTGGCCCCATG CCCTGGCAAT
15901 CTGAATGAGG CCAAATTTA AAGGAGTGG AACTACGAT TGTCAAGAGAA AATATCAAGA
15961 CAGACCACCA CTCAGGCTAT GCGGTGTTG TGACCGACCA GCTACTCTTA GCCAGCTCTA
16021 TTGTGAAATT CCAGAGCAAT TATCAGAGCA TGAAGATAAC TACAGTTAG TGAAGTAAGG
16081 GGTGTGGGTC CCTAAGTGGA TGGTGCATAA ATCTATGTAG GTGATGCCCTA AGTGCACATT
16141 GATAATCCAA AATATCAGCA ATGTGGAATG TCTTCCAAGG AGACACTGTAG ACACACATT
16201 TAGAACTTTG CTCATGGCTC TAATAAATAG CTAGCTAGAA ATCATTTCCCT GAAGAGGTTA
16261 GTCTGAGTTA CGGTTCCAGG GCAAACATTC AGTGTGGCA AGGAAGGGCAT TGCACTCAGG
16321 AGCCAAAGGT CAGCTGGTCA CATTGCTACAGA AGAGTAGAGA GTCAAGAGTGT GAGTAGAAAG
16381 AGGATACAGG TTATAAAACCTCACTGTCACTCCTCAGCAA TCCATTTCT CCTAAAAGGC
16441 TTTACCTTCT AAAGATTAA GTCTTCAAA CCAGTACCGAG TAGCCTGGGA ACAAAAGTTG
16501 AAACAAATGA GCCTTGTGG GGCATTTCAC ACTTAAACAA GGGCATCACC TAGGAGGAGC
16561 CCTGTGTGCA GTAGGAAGTG TGGCCTCTGT GTCAAGGATG CTCAGGCTAA TAAGGGTCC
16621 TCTATCTGAG GGACCCATG AAGATTCAAC AAGTAGTTGT GAGAATTCCC TGAAATGGA
16681 TGCTACCAAT TTGACATTG TAGACCTGCT ATTGTGTGCT TCTTATTGG GCTCTCCCAT
16741 CTCCCAACTT TCCAACCCAT ATTCCACATT AATCCCTTCC ACCACCATGC AACACTAGGT
16801 AGGAGAGAAG GAAGGTTAGA AGAGAAAGTG GGTATAGATC TATTTAGACT ACTTCCTGCT
16861 GATTAGGGC AAGTCCAATC GTCAATTGTCAGGATACCTCC ACCAGCAAC CAGCAAACCA
16921 GCAAATCAGA AACAGCAAA CGAGCCAACA AGGCAGCACT ACCAGCAGG ATTGGGGTCG
16981 GTAGCGTGGG AGCAGTCACT ACTGGTCTTC TCATGGCTTT GGCAATACTA CTCTCTCAAG
17041 AAATTCCGTA ATTTTTTCCC CACCACTGA AATTCCGTA TTTAAATGC AAAACTATCTA
17101 CAGCTGGCAA AAATCACATC TCTCCTAGAG CACAAGACAA ATCATAGTTA CTGGCTATT
17161 GCAATCTGAA GCATCTCAAT ATCCCACACC TGGGATTAAA ACAAAACAT ATTCACATCA
17221 CATAACTGTT TTTTTTCCC AATTTTTAT TAGGTATTT CTTTATTAC ATTTCAAATG
17281 CTATCCCGAA AGTCCCTAT ACCCTCCCCAC CTCCCTGCTC CCCTACACAC CCAACTCCAC

Figure 5G:

17341 TTTTGACCC TGGAGTTCCC CGGTACTGGG GCATATAAAG TTTGCAAGAC CAAGGGGCCT
 17401 CTCTCCCAG TGATGGCCGA CTAAGCCATC TTCTGCTACA TATGCAGATA GAGACACGAG
 17461 CTCTGGGGT ACTAGTTAGT TCATATTGTT GTTCCACCTA TAGGGTCGCA GACCCCTTCA
 17521 GCTCCTTGGG TACTTTGTCT AGCTCCTCCA CTGGGGCTC TGTGTTTAT CTAATAGATG
 17581 ACTGTGAGCA TCCACTTCTG TATTTGACAG GCACTGGCCT AGCGTCACAT GAGCCAGCTA
 17641 TATCAGGGTC CTTTCAGCAA AACCTTGCTG GCATGTGCAA TAGTGTCTGC GTTTGGTGGT
 17701 TGATTATGGG ATGGATCCAC TAGTTCTAGA CGGGCCGCA CGCGGGTGA GCTCCAGCTT
 17761 TTGTTCCCTT TAGTGAGGGT TAATTGCGC CTTGGCGTAA TCATGGTCAT AGCTGTTCC
 17821 TGTGTGAAT TGTATTCGC TCACAATTCC ACACAAACATA CGAGCGGAA GCATAAAAGTG
 17881 TAAAGCCTGG GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTCGCTTGC GCTCACTGCC
 17941 CGCTTCCAG TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGG
 18001 GAGAGGGGGT TTGCGTATTG GGCGCTCTC CGCTTCTCG CTCACTGACT CGCTGCGCTC
 18061 GGTGTTCCGG CTGCGGCGAG CGGTATCAGC TCACTCAAAG GCGGTAAATAC GGTTATCCAC
 18121 AGAATCAGGG GATAACGCAG GAAAGAACAT GTGAGCAAA GGCAGCAAA AGGCCAGGAA
 18181 CCGTAAAAG GCCGCGTTGC TGGCGTTTT CCATAGGCTC CGCCCCCTG ACGAGCATCA
 18241 CAAAAATCGA CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGAATATAAA GATACCAGGC
 18301 GTTTCCTCCCT GGAAGCTCCC TCGTGCCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA
 18361 CCTGTCGCCCTT CCGGAAGCGT GGCCTTCT CATAGCTCAC GCTGTAGGTA
 18421 TCTCAGTTCG GTGTAGGTG TCGCCTCCAA GCTGGCTGT GTGCACGAAC CCCCCGTTCA
 18481 GCCCGACCCGC TCGCGCTTAT CCGGTTAATC TCGTCTTGAG TCCAACCCGG TAAGACACGA
 18541 CTTATCGCCA CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGGG
 18601 TGCTACAGAG TTCTTGAAGT GGTGGCCTAA CTACGGCTAC ACTAGAAGGA CAGTATTGG
 18661 TATCTGCGCT CTGCTGAAGC CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG
 18721 CAAACAAACC ACCGCTGGTA GCGGTGGTTT TTTTGGTTGC AAGCAGCAGA TTACGCGCAG
 18781 AAAAAAAGGA TCTCAAGAAG ATCCTTTGAT CTTTTCTACG GGGTCTGAGC CTCAGTGGAA
 18841 CGAAAACCTCA CGTTAAGGGG TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT
 18901 CCTTTAAAT TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT AAACCTGGTC
 18961 TGACAGTTAC CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTGGTTC
 19021 ATCCATAGTT GCCTGACTCC CGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC
 19081 TGGCCCGAGT GCTGAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC
 19141 AATAAACCGAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCTC
 19201 CATCCAGTCT ATTAAATTGTT GCCGGGAAGC TAGAGTAAGT AGTTGCCAG TTAATAGTT
 19261 GCGCAACGTT GTGCCATTG CTACAGGCAT CGTGGTGTCA CGCTCGTGT TTGGTATGGC
 19321 TTCATTAGC TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCA TGGTGTGCAA
 19381 AAAAGCGGTT AGCTCCTTCG GTCCCTCGAT CGTTGTCAGA AGTAAGTTGG CCGCAGTGGT
 19441 ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG
 19501 CTTTTCTGTG ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTG TGCGGCGACC
 19561 GAGTTGCTCT TGCCCGCGGT CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA
 19621 AGTGCTCATC ATTGGAAAAC GTTCTTCGGG GCGAAAACCTC TCAGGATCT TACCGCTGTT
 19681 GAGATCCAGT TCGATGTAAC CCACTCGTGA ACCCAACTGA TCTTCAGCAT CTTTACTTT
 19741 CACCAAGCGT TCTGGGTGAG CAAAAAACAGG AAGGCAAAAT GCGCAAAAAA AGGGATAAG
 19801 GGCACACGG AAATGTTGAA TACTCATACT CTTCTTTTT CAATATTATT GAAGCATTAA
 19861 TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT
 19921 AGGGGTTCCG CGCACATTTC CCCGAAAAGT GCCACCTAA TTGTAAGCGT TAATATTTG
 19981 TTAACATTG CGTTAAATT TGTTAAATC AGCTCATTTC TTAACCAATA GGCGAAATC
 20041 GGAAAATCC CTTATAAAATC AAAAGAATAG ACCGAGATAG GTTGGAGTGT TGTTCCAGTT
 20101 TGGAACAGA GTCCACTATT AAAGAACGTG GACTCCAACG TCAAAGGGCG AAAAACCGTC
 20161 TATCAGGGCG ATGGCCCCT ACGTGAACCA TCACCTTAAT CAAGTTTTT GGGGTGAGG
 20221 TGCGTAAAG CACTAAATCG GAACCTAAAGG GGGAGCCCCC GATTTAGAGC TTGACGGGG
 20281 AAGCCGGCGA ACGTGGCGAG AAAGGAAGGG AAGAAAGCGA AAGGAGCGGG CGCTAGGGCG
 20341 CTGGCAAGTG TAGCGGTAC GCTGCGCTA ACCACCCAC CGGCCGCGCT TAATGCGCCG
 20401 CTACAGGGCG CGTCCCATTC GCCATTCAAG CTGCGCAACT GTTGGGAAGG GCGATCGGTG
 20461 CGGGCCTCTT CGCTATTACG CCAGCTGGCG AAAGGGGGAT GTGCTGCAAG GCGATTAAGT
 20521 TGGGTAACGC CAGGGTTTC CGAGTCACGA CGTTGTAAAA CGACGGCCAG TGAGCGCG
 20581 TAATACGACT CACTATAGGG CGAATTGGGT ACCGGGCCCC CCC

//

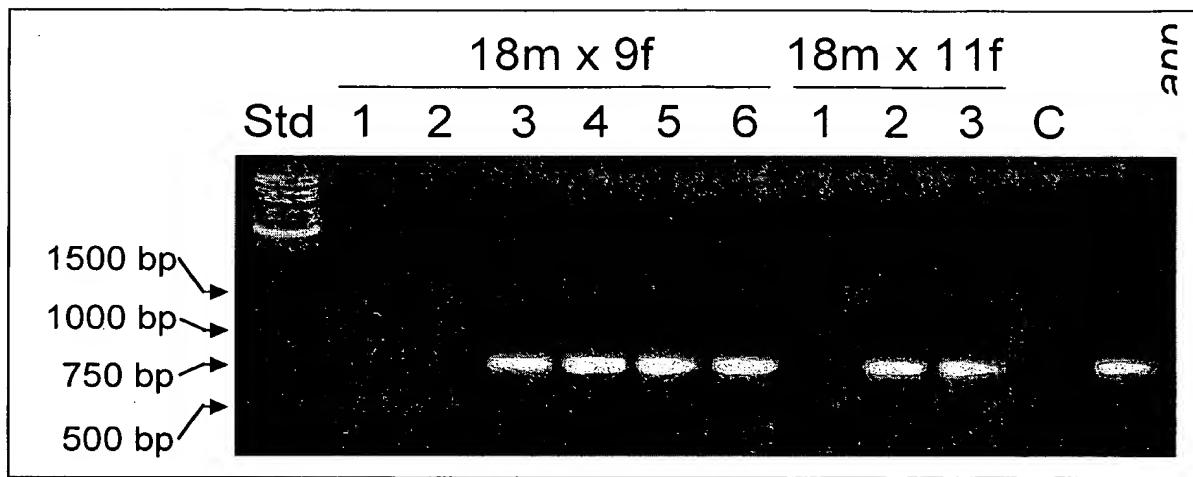


Figure 6

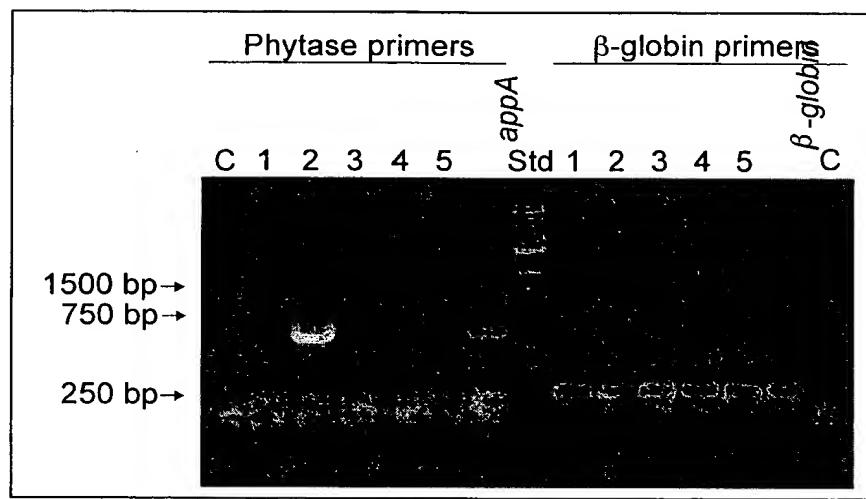


Figure 7

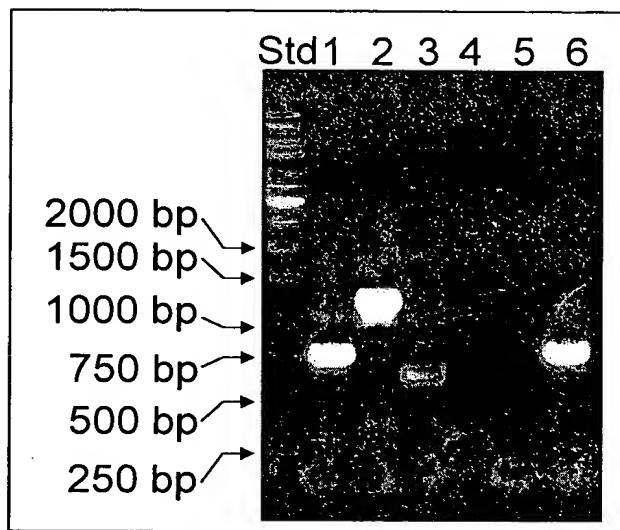


Figure 8

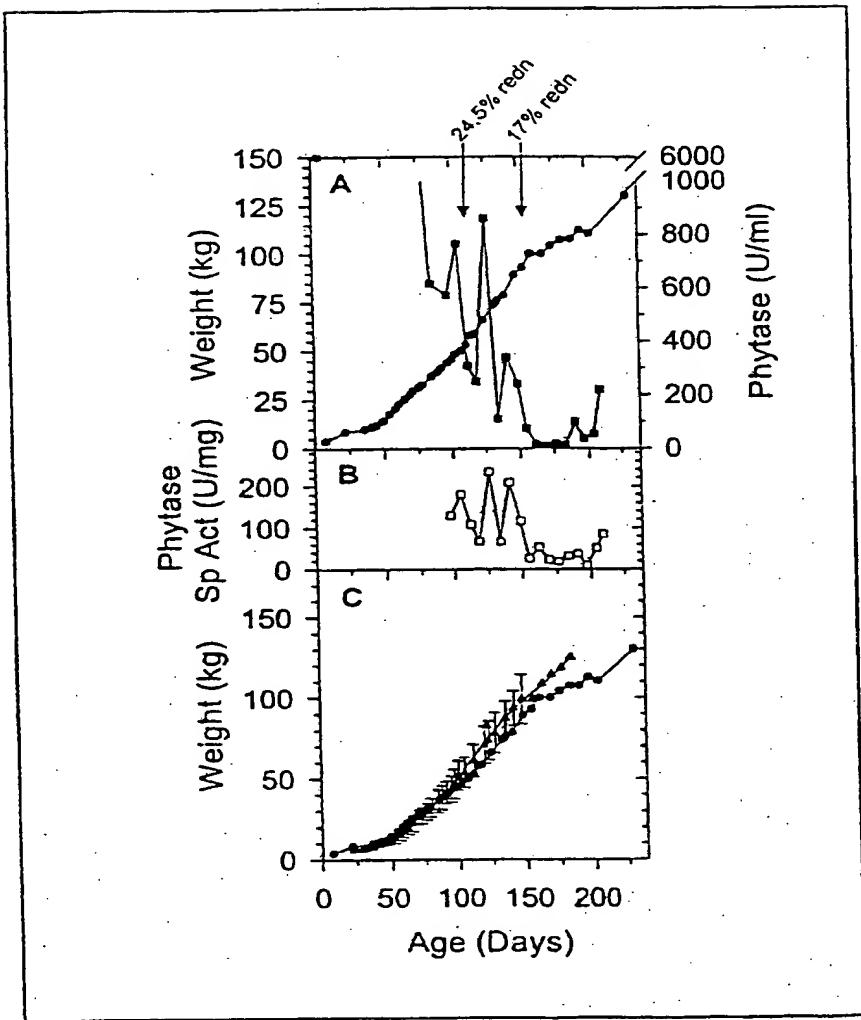


Figure 9

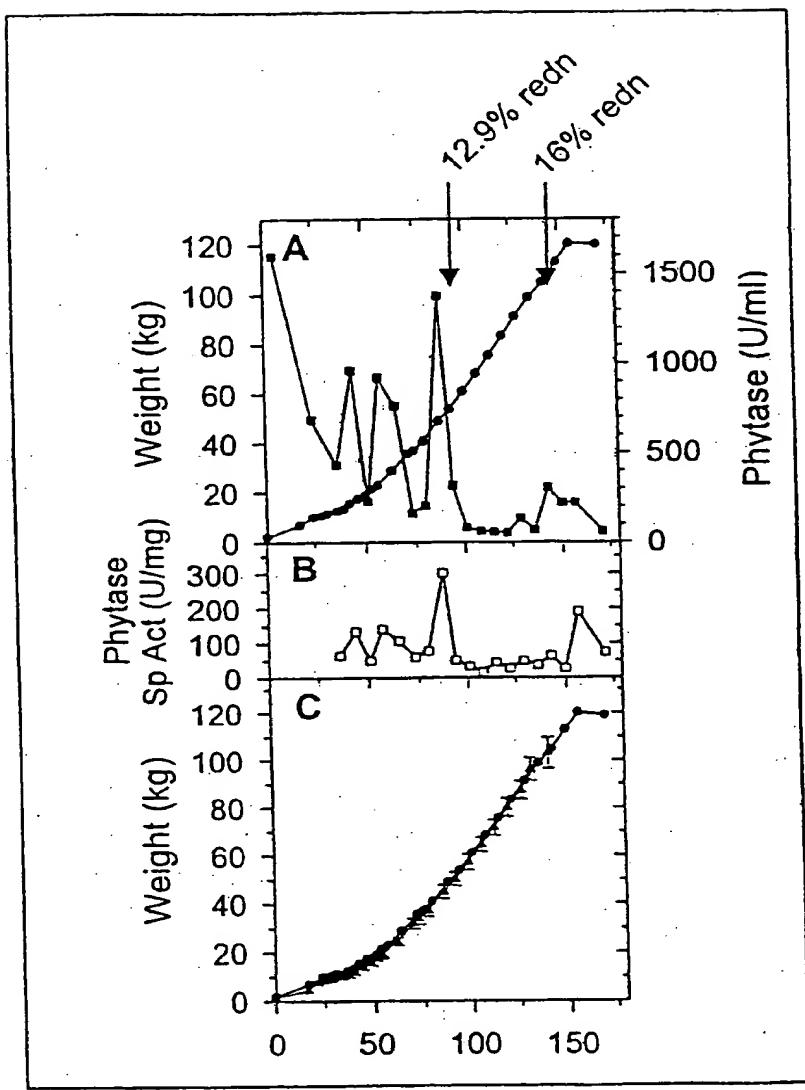


Figure 10

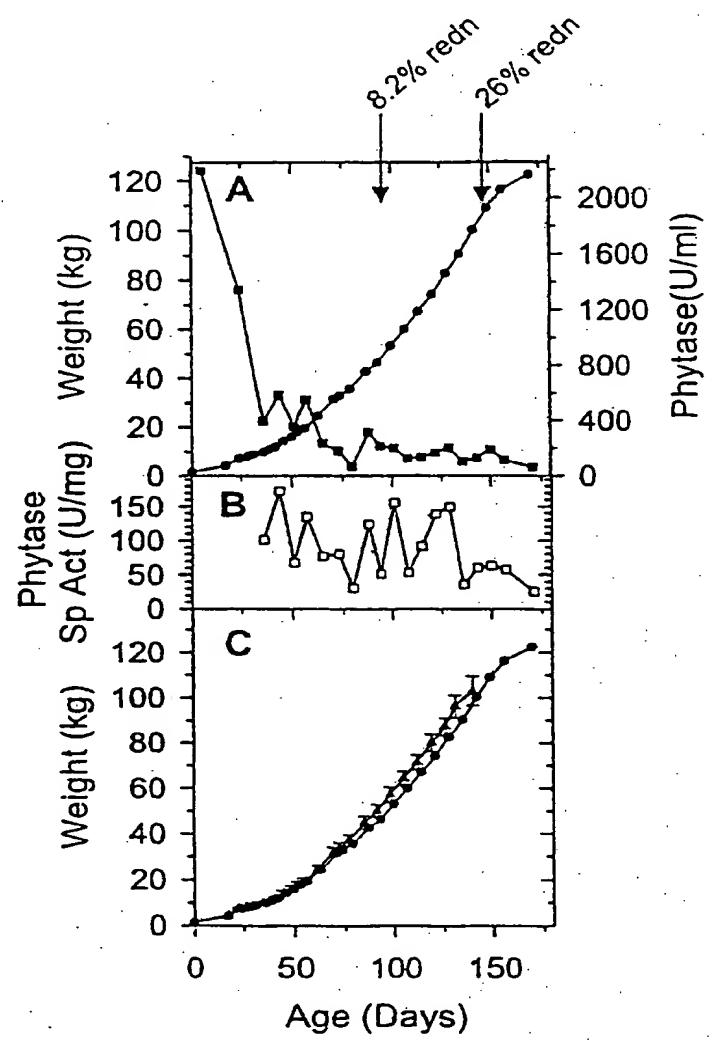


Figure 11

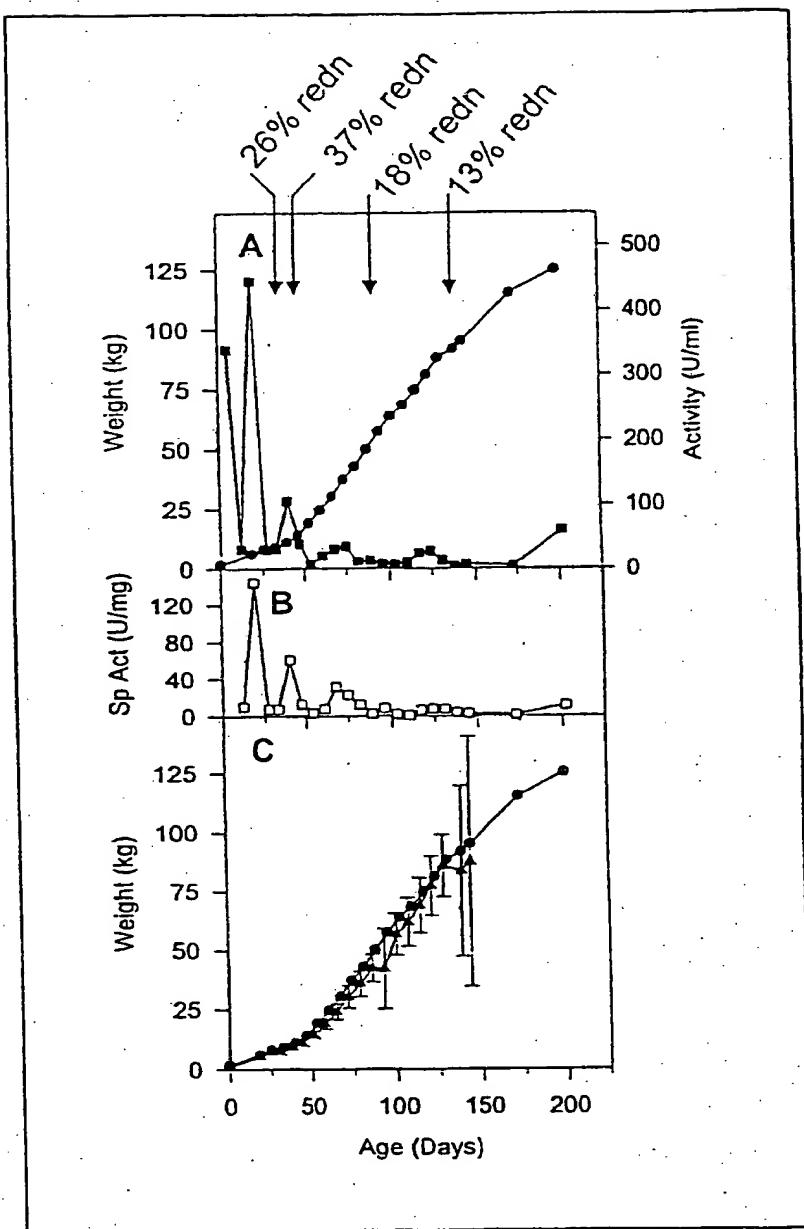


Figure 12

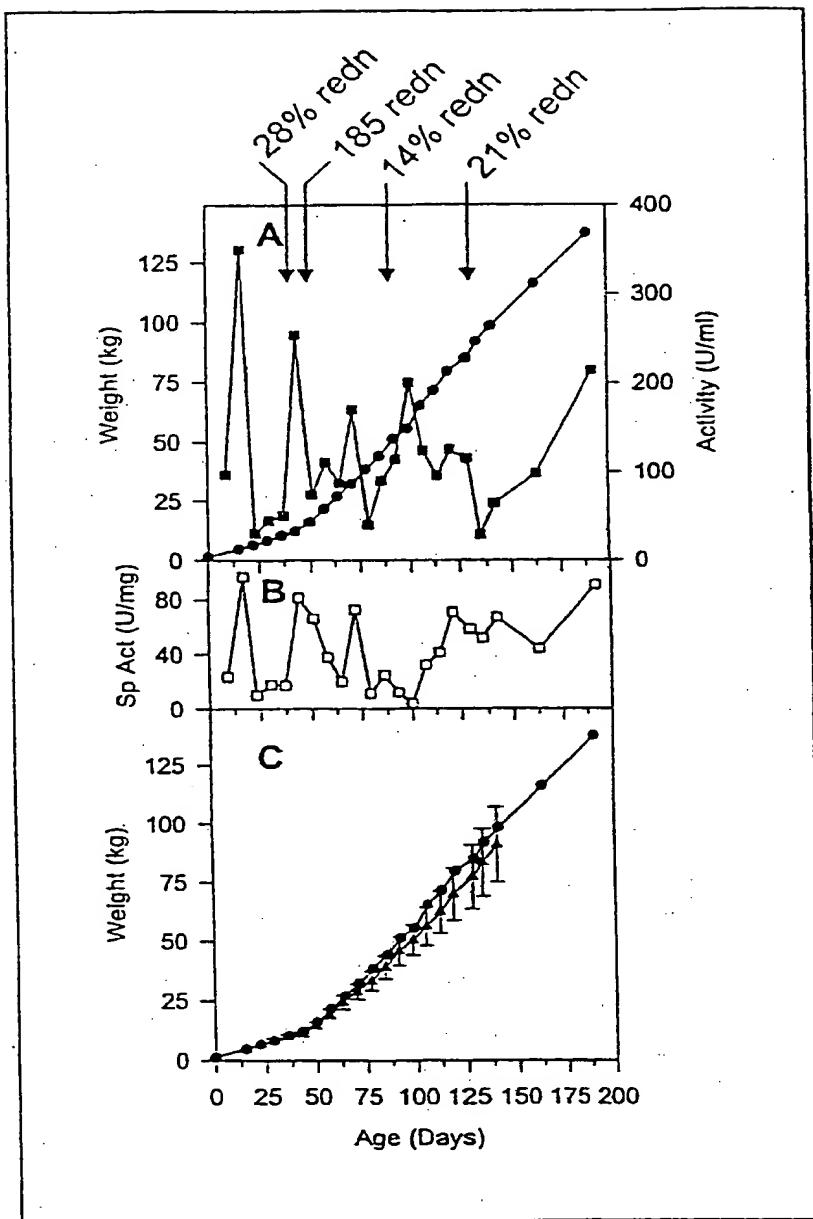


Figure 13

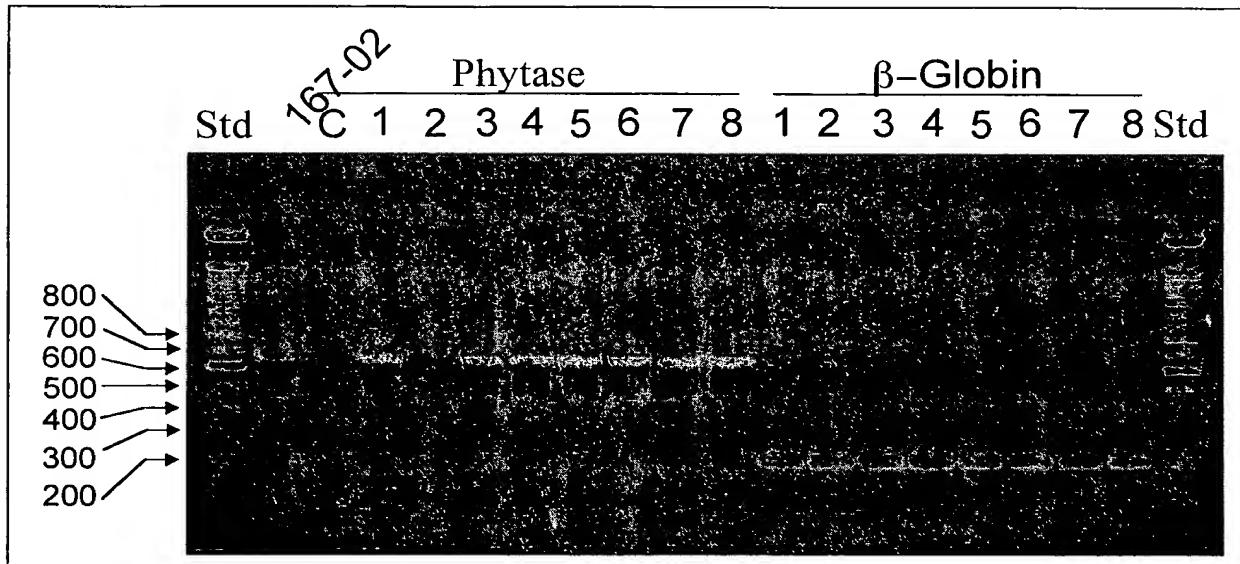


Figure 14

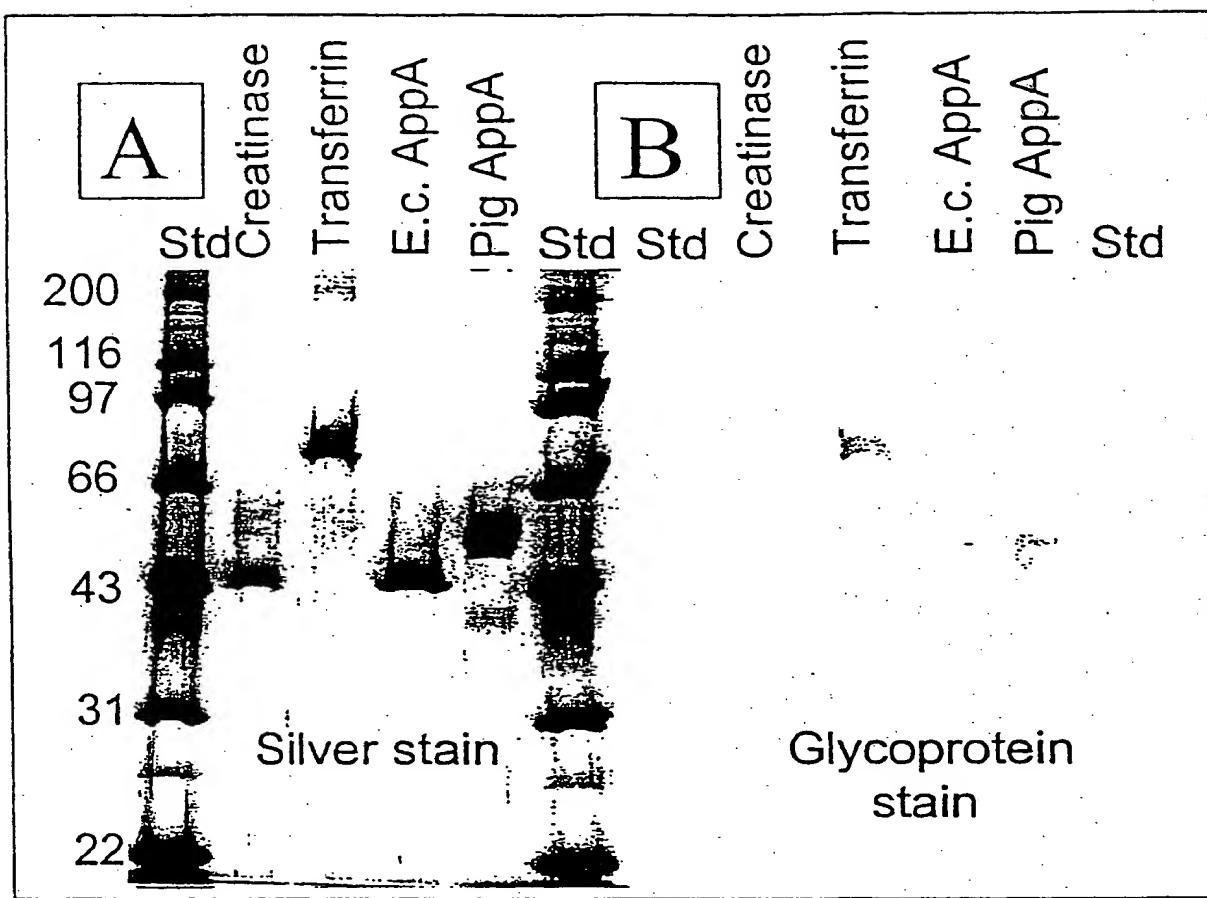


Figure 15

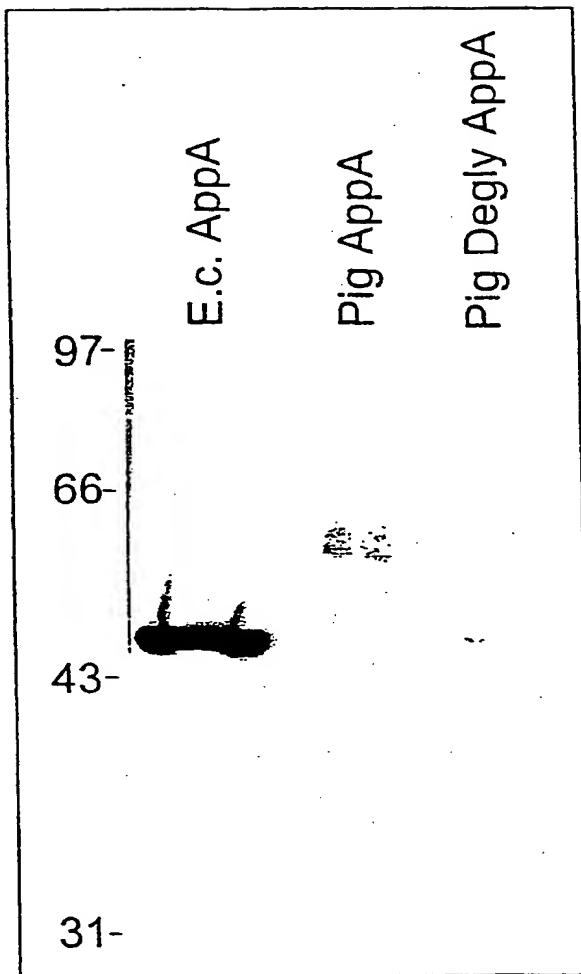


Figure 15B

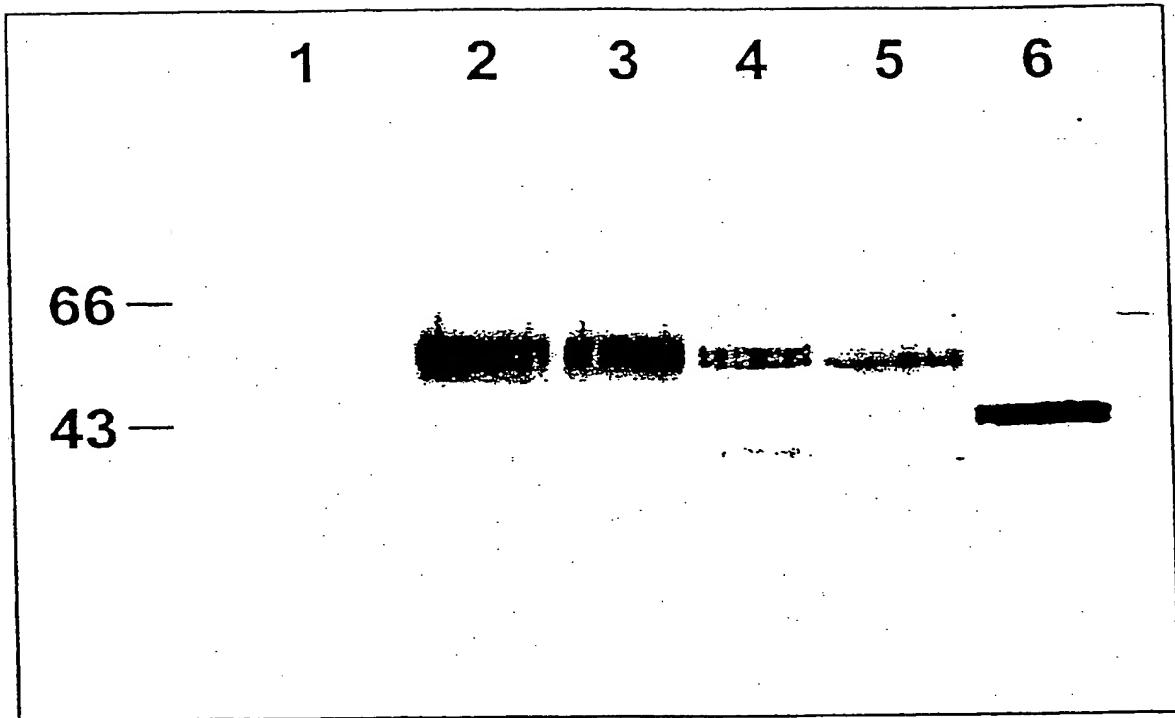


Figure 16

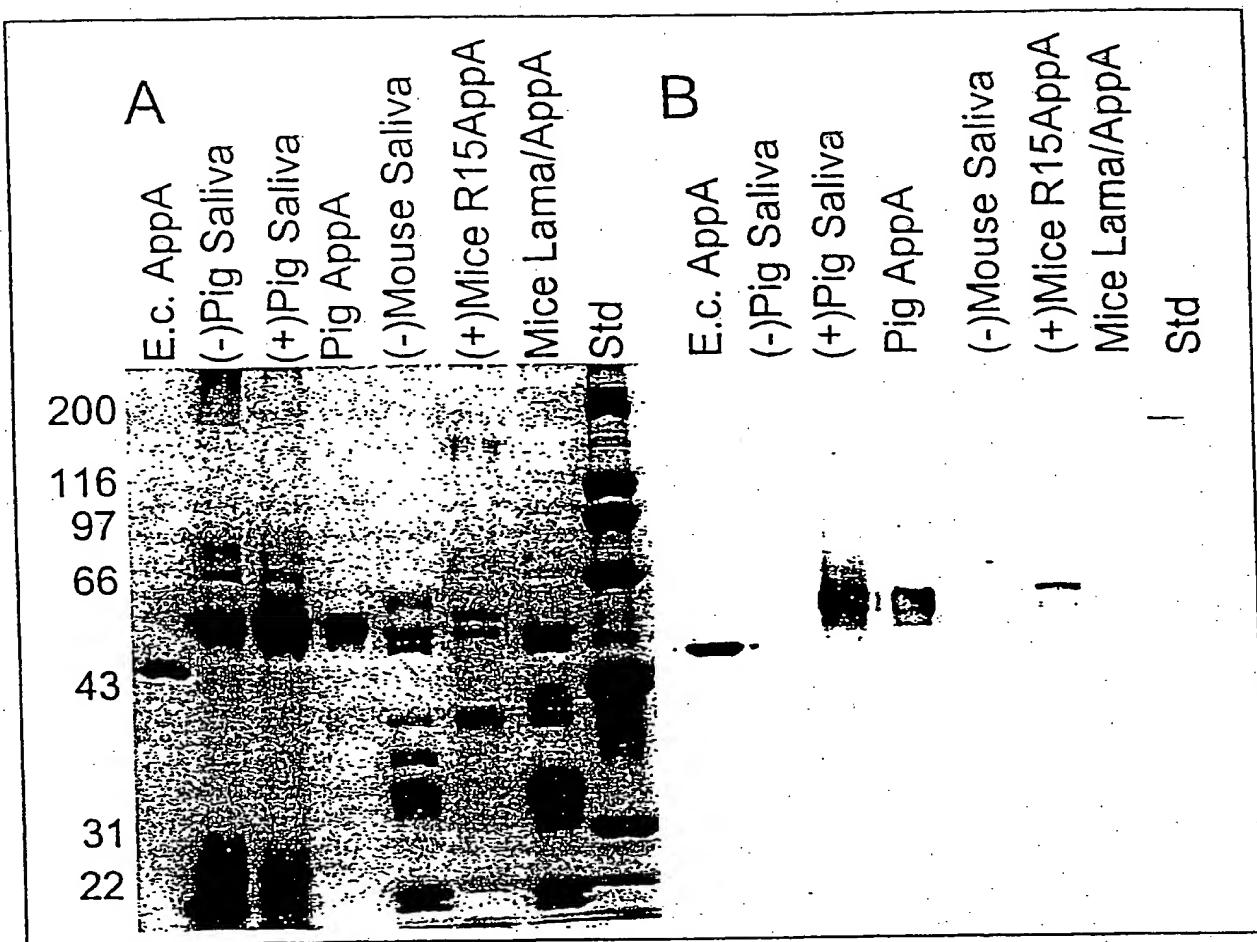


Figure 17

Figure 18: Nucleic acid sequence of the known segment of the R15/appa+intron plasmid, including the vector sequences of pBLCAT3 (SEQ ID NO:2).

LOCUS R15/appa+intron 6708 bp DNA SYN 15-APR-2000
DEFINITION R15/appa+intron transgene with vector cut 13543 to 4954
ACCESSION R15/appa+intron
REFERENCE 1 (bases 1 to 6708))
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA gene; periplasmic phosphoanhydride phosphohydrolase; artificial sequence;
AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
ACCESSION M64793 M36414
VERSION M64793.1 GI:206711
SOURCE Rat (Sprague-Dawley) liver DNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1748)
AUTHORS Lin, H.H. and Ann, D.K.
TITLE Molecular characterization of rat multigene family
encoding proline-rich proteins
JOURNAL Genomics 10, 102-113 (1991)
MEDLINE 91257817
FEATURES Location/Qualifiers
source 1..1748
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="liver"
/tissue_lib="cosmid genomic library"
misc_feature 1802-1810
/function=" consensus sequence for initiation in
higher eukaryotes "
FEATURES Location/Qualifiers
DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,
ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
VERSION M58708.1 GI:145283
SOURCE Escherichia coli DNA.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision;
Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1811..3109)
AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

Figure 18A:

TITLE The complete nucleotide sequence of the *Escherichia coli* gene appA reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase
JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE 90368616

FEATURES Location/Qualifiers

Source 1811..3109
/organism="Escherichia coli"
/db_xref="taxon:562"
sig_peptide 1811.. 1876
/gene="appA"
CDS 1811..3109
/gene="appA"
/standard_name="acid phosphatase/phytase"
/transl_table=11
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AAA72086.1"
/db_xref="GI:145285"
mat_peptide 1877 3106
/gene="appA"
/product="periplasmic phosphoanhydride phosphohydrolase"
mutation replace(1817.. 1819, "gcg changed to gcc")
/gene="appA"
/standard_name="A3 mutant"
/note="created by site directed mutagenesis"
/phenotype="silent mutation"
mutation replace(3092..3094, " ccg changed to ccc")
/gene="appA"
/standard_name=" P428 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "
mutation replace(3095..3097, " gcg changed to gct")
/gene="appA"
/standard_name=" A429 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "

Figure 18B:

DEFINITION Plasmid pBLCAT3 (bases 3109 to 6708)
ACCESSION X64409
VERSION X64409.1 GI:58163
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 3109 to 6708)
AUTHORS Luckow, B.H.R.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG
REFERENCE 2 (bases 3109 to 6708)
AUTHORS Luckow, B. and Schutz, G.
TITLE CAT constructions with multiple unique restriction sites
for the functional analysis of eukaryotic promoters and
regulatory elements
JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)
MEDLINE 87260024
COMMENT Promoterless CAT vector for transient transfection
experiments with eukaryotic cells. Allows the analysis of foreign
promoters and enhancers.
FEATURES Location/Qualifiers
source 3109 to 6116
/organism="synthetic construct"
/db_xref="taxon:32630"
SV40 t intron 3197..3810
/note="SV40 signals"
polyA_signal 3807..4047
/note="SV40 signals"
CDS complement(5244..6104)
/codon_start=1
/transl_table=11
/gene="Amp"
/product="beta-lactamase"
/protein_id="CAA45753.1"
/db_xref="GI:58165"
BASE COUNT 1916 a 1479 c 1515 g 1798 t
ORIGIN
1 GGATCCCCTT TGCTATGTAG TTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGGTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTTGTTT CTAGCATAAC CAAAAGATT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCGAA AATATTACCA TGATACTGAG CATTGTAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTAA CATGATTTTC ATTAATCAGG
301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
361 TATTTCACTA AACTAGTTT ATCTATTTG TTGCTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAAT TTTCTACATA
481 TATCCTGGTT AGAGAGTGCT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
601 TGGGAAGAAA CCATTGGTG ACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG
661 TAAAACATAT GTTGACCAAG CCCTTCTTTT CAATAGGCTT AATGTGAATA AAATGTAAA
721 GATTCTCTTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT

Figure 18C:

901 TAAGATAAAAG GTAACTGTAT ACATTGTC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATTCTTGC CTTAAACAAAC TTAAATAGAA TGATTAAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCCT
1081 GATACTAACCA CAGGTAATTC CCACACGTGT TTTCACACTA CATTTGCTGG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTGCGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTAACG TGACTATTG ATCAAAGAAA TTTATTACCT TCAGTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAACCTCA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTAA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTCTTA CACATGATT TATTCAAATT TACTCTAAC
1561 ATTGTTAACG CATTAGAAA AGGCATACTG GCAACTTTC CTTACCTCAT CCAGCTGGC
1621 AAAAGTCCCAGTGTGGAGTA AAGGATGCAA GATTTCCTGC TCTGTTAAGT ATAAAATAAT
1681 AGTATGAATT CAAAGGTGCC ATTCTCTGC TTCTAGTTA AAAGGCAGTG CTTGCTTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTG GCGAGATTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCACC ATGAAAGCCA TCTTAATCCC ATTTTATCT CTTCTGATTC CGTTAACCCC
1861 GCAATCTGCA TTGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCAGA
1981 CGCATGGCCA ACCTGGCCGG TAAAATGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCAAAAAA
2101 GGGCTGCCCG CAGTCTGGTC AGGTCGCGAT TATTGCTGAT GTGACGAGC GTACCCGTA
2161 AACAGGCAGA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCA
2221 GGCAGATACG TCCAGTCCCC ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
2281 GGATAACGCG AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACTTGTGC CTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
2461 ATCGGAACTC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTC TCCTGCAACA AGCACAGGGAGTGGCAGGAGC CGGGGTGGGG
2581 AAGGATCACC GATTCAACACC AGTGGAACAC CTTGCTAAGT TTGCATAACG CGCAATTAA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCC ACCCCGTTAT TAGATTGAT
2701 CAAGACAGCG TTGACGCCCG ATCCACCGCA AAAACAGGGC TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATGCCCG GACACGATAC TAATCTGGCA AATCTGGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCCG GTCAAGCCGTA AACACGCCG CCAGGTGGTGT AACTGGTGT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGGTCTTCCA
2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAAG GTATAAAGCA
3121 GTTATTGGTG CCCTTAAACG CCTGGTCTA CGCCTGAATA AGTATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTTGTTGTG ACATAATTGG
3241 ACAAACTACC TACAGAGATT TAAAGCTCTA AGGTAATAT AAAATTAAAT AGTGTATAAT
3301 GTGTTAAACT ACTGATTCTA ATTGTTGTG TATTTTAGAT TCCAACTTAT GGAACGTGATG
3361 AATGGGAGCA GTGGTGAAT GCCTTAAATG AGGAAAACCT GTTTGCTCA GAAGAAATGC
3421 CATCTAGTGA TGATGAGGCT ACTGCTGACT CTCAACATTC TACTCCTCCA AAAAAGAAGA
3481 GAAAGGTAGA AGACCCCAAG GACTTCCCT CAGAATTGCT AAGTTTTTG AGTCATGCTG
3541 TGTTTAGTAA TAGAACTCTT GCTTGCTTGT CTATTACAC CACAAAGGAA AAAGCTGCAC
3601 TGCTATACAA GAAAATTATG GAAAATATT CTGTAACCTT TATAAGTAGG CATAACAGTT
3661 ATAATCATAA CATACTGTTT TTTCTTACTC CACACAGGCC TAGAGTGTCT GCTATTAATA
3721 ACTATGCTCA AAAATTGTGT ACCTTAGCT TTTTAATTG TAAAGGGTT AATAAGGAAT
3781 ATTTGATGTA TAGTGCCTTG ACTAGAGATC ATAATCAGCC ATACCACATT TGTAGAGGTT
3841 TTACTTGCTT TAAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA AATGAATGCA
3901 ATTGTTGTTG TTAACCTGTT TATTGCTGCT TATAATGGTT ACAAAATAAG CAATAGCATC
3961 ACAAAATTCA CAAATAAACG ATTTTTCTA CTGCATTCTA GTTGTGGTT GTCCAAACTC
4021 ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCAGGTAC CGAGCTCGAA TTCGTAATCA
4081 TGGTCATAGC TGTTCCTGT GTGAAATTGT TATCCGCTCA CAATTCCACA CAACATACGA
4141 GCCGGAAGCA TAAAGTGTAA AGCCTGGGT GCCTAATGAG TGAGCTAATC CACATTAATT
4201 GCGTTGCGCT CACTGCCGC TTTCCAGTCG GGAAACCTGT CGTGCAGCT GCATTAATGA
4261 ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGC GCTCTCCGC TTCCTCGCTC
4321 ACTGACTCGC TGCCTCGGT CGTTGGCTG CGGCGAGCGG TATCAGCTCA CTCAAAGGCG

Figure 18D:

4381 GTAATACGGT TATCCACAGA ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC
4441 CAGCAAAAGG CCAGGAACCG TAAAAAAGGCC GCGTTGCTGG CGTTTTCCA TAGGCTCCGC
4501 CCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA
4561 CTATAAAGAT ACCAGGCGTT TCCCCCTCGA AGCTCCCTCG TGCGCTCTCC TGTTCCGACC
4621 CTGCCGCTTA CCGGATACCT GTCCGCCCTT CTCCCTCGG GAAGCGTGGC GCTTTCTCAA
4681 TGCTCACGCT GTAGGTATCT CAGTTCCGGTG TAGGTCGTTG GCTCCAAGCT GGGCTGTGTG
4741 CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAATATCG TCTTGAGTCC
4801 AACCCGGTAA GACACGACTT ATGCCCACTG GCAGCAGCCA CTGTTAACAG GATTAGCAGA
4861 GCGAGGTATG TAGGCCGTGC TACAGAGTTC TTGAAGTGGT GGCTTAACTA CGGCTACACT
4921 AGAAGGACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTTCGG AAAAAGAGTT
4981 GGTAGCTCTT GATCCGGCAA ACAAAACCACC GCTGGTAGCG GTGGTTTTTG TGTTGCAAG
5041 CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
5101 TCTGACGCTC AGTGGAACGA AAACTCACGT TAAGGGATTT TGTCATGAG ATTATCAAAA
5161 AGGATCTTCA CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT CTAAAGTATA
5221 TATGAGTAAA CTTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC TATCTCAGCG
5281 ATCTGTCTAT TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA
5341 CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG
5401 GCTCCAGATT TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGCCT
5461 GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG AGTAAGTAGT
5521 TCGCCAGTTA ATAGTTGCG CAACGTTGTT GCCATTGCTA CAGGCATCGT GGTGTCACGC
5581 TCGTCGTTG GTATGGCTTC ATTCAAGCTCC GGTTCCCAAC GATCAAGGCG AGTTACATGA
5641 TCCCCCATGT TGTGCAAAAA AGCGGTTAGC TCCTTCGGTC CTCCGATCGT TGTCAGAAGT
5701 AAGTTGGCCG CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCTAATTC TCTTACTGTC
5761 ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA
5821 TAGTGTATGC GGCAGCGAG TTGCTCTTGC CCGCGTCAA TACGGGATAA TACCGGCCA
5881 CATAGCAGAA CTTAAAAGT GCTCATCATT GGAAACCGTT CTTCGGGGCG AAAACTCTCA
5941 AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAACCCA CTCGTGCACC CAACTGATCT
6001 TCAGCATCTT TTACTTTAC CAGCGTTCT GGGTGAGCAA AAACAGGAAG GCAAAATGCC
6061 GCAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT CCTTTTCAA
6121 TATTATTGAA GCATTATCA GGGTTATTGT CTCATGAGCG GATACATATT TGAATGTATT
6181 TAGAAAAATA AACAAATAGG GGTTCCCGC ACATTTCCCC GAAAAGTGCC ACCTGACGTC
6241 TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC GAGGCCCTT
6301 CGTCTCGCGC GTTCCGGTGA TGACGGTGAA AACCTCTGAC ACATGCAGCT CCCGGAGACG
6361 GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG CGCGTCAGCG
6421 GGTGTTGGCG GGTGTCGGGG CTGGCTTAAC TATGCGGCAT CAGAGCAGAT TGTACTGAGA
6481 GTGCACCATATA TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAATA CCGCATCAGG
6541 CGCCATTGCGC CATTCAAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG GGCCTCTCG
6601 CTATTACGCC AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG GGTAACGCCA
6661 GGGTTTCCC AGTCACGAGC TTGTAAAACG ACGGCCAGTG CCAAGCTT
//

Figure 19: Nucleic acid sequence of the known segment of the R15/appa+intron transgene used for the generation of transgenic mice (SEQ ID NO: 3).

LOCUS R15/appa 4060 bp DNA SYN 15-APR-2000
DEFINITION R15/appa transgene without vector
ACCESSION R15/appa
REFERENCE 1 (bases 1 to 4060)
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA gene; periplasmic phosphoanhydride phosphohydrolase; artificial sequence;
AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
ACCESSION M64793 M36414
VERSION M64793.1 GI:206711
SOURCE Rat (Sprague-Dawley) liver DNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1748)
AUTHORS Lin, H.H. and Ann, D.K.
TITLE Molecular characterization of rat multigene family
encoding proline-rich proteins
JOURNAL Genomics 10, 102-113 (1991)
MEDLINE 91257817
FEATURES Location/Qualifiers
source 1..1748
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="liver"
/tissue_lib="cosmid genomic library"
misc_feature 1802-1810
/function=" consensus sequence for initiation in
higher eukaryotes "
FEATURES Location/Qualifiers
DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA) gene,
ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
VERSION M58708.1 GI:145283
SOURCE Escherichia coli DNA.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision;
Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 1811..3109)
AUTHORS Dassa, J., Marck, C. and Boquet, P.L.

Figure 19A:

TITLE The complete nucleotide sequence of the *Escherichia coli* gene *appA* reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase
JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE 90368616

FEATURES Location/Qualifiers

Source 1811..3109
/organism="Escherichia coli"
/db_xref="taxon:562"
sig_peptide 1811..1876
/gene="appA"
CDS 1811..3109
/gene="appA"
/standard_name="acid phosphatase/phytase"
/transl_table=11
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AAA72086.1"
/db_xref="GI:145285"
mat_peptide 1877 3106
/gene="appA"
/product="periplasmic phosphoanhydride phosphohydrolase"
mutation replace(1817.. 1819, "gcg changed to gcc")
/gene="appA"
/standard_name="A3 mutant"
/note="created by site directed mutagenesis"
/phenotype="silent mutation"
mutation replace(3092..3094, " ccg changed to ccc")
/gene="appA"
/standard_name=" P428 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "
mutation replace(3095..3097, " gcg changed to gct")
/gene="appA"
/standard_name=" A429 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "

Figure 19B:

```
SV40 t intron      3197..3810
                  /note="SV40 signals"
polyA_signal      3807..4047
                  /note="SV40 signals"

BASE COUNT      1257 a      814 c      843 g      1146 t
ORIGIN

 1 GGATCCCCTT TGCTATGTAG TTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
 61 GAGAGTCCTG TTTGGTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTTGTT CTAGCATAAC CAAAAGATT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCCGAA AATATTACCA TGATACTGAG CATTGTAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTAA CATGATTTTC ATTAATCAGG
301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGCACA
361 TATTTCACTA AACTAGGTT ATCTATTTG TTGCTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
481 TATCCTGGTT AGAGAGTGT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
601 TGGGAAGAAA CCATTGGTG ACAAATATT CAAATAAAA TAGACAAACA TAGTTAATTG
661 TAAAACATAT GTTGACCAAG CCCTTCTTT CAATAGGTT AATGTGAATA AAATGTAAA
721 GATTCTCTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
901 TAAGATAAAAG GTAACTGTAT ACATTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATTCTTGC CTTAAACAAC TTAAATAGAA TGATTAAAA TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCTT
1081 GATACTAACCA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTGCTGG GATTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTGGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTAACG TGACTATTG ATCAAAGAAA TTTATTACCT TCAGTTCAA TGGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAACCTTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTAA TCAATCAATT
1501 GTATGTATCA ATATATGGC TATTTCTTA CACATGATT TATTCAAATT TACTCTAAC
1561 ATTGTTAAC CATTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGC
1621 AAAAGTCCC GTGTGGAGTA AAGGATGCAA GATTTCCTGC TCTGTTAAGT ATAAAATAAT
1681 AGTATGAATT CAAAGGTGCC ATTCTCTGC TTCTAGTTAT AAAGGCAGTG CTTGCTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCT CTTCTGATTC CGTTAACCC
1861 GCAATCTGCA TTGCTCAGA GTGAGCCGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCAGA
1981 CGCATGGCCA ACCTGGCCGG TAAAACCTGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCAAAAAA
2101 GGGCTGCCCG CAGTCTGGTC AGGTGCGAT TATTGCTGAT GTGACGAGC GTACCCGTA
2161 AACAGGCAGA GCCTTCGCCG CGGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCC
2221 GGCAGATACG TCCAGTCCCC ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
2281 GGATAACCGC AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTGCAGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACTTGTGC CTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTAC
2461 ATCGGAACTC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTG TCCTGCAACA AGCACAGGGG ATGCCGGAGC CGGGGTGGGG
2581 AAGGATCACC GATTACACCC AGTGGAAACAC CTTGCTAAGT TTGCTAAACG CGCAATTAA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCGCC ACCCCGTTAT TAGATTGAT
2701 CAAGACAGCG TTGACGCCCG ATCCACCGCA AAAACAGGGC TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTGGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCCG GTCAGGCCGA TAACACGCCG CCAGGTGGTG AACTGGTGT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGGTCTCCA
```

Figure 19C:

2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CCGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAAG GTATAAGGC
3121 GTTATTGCTG CCCTTAAACG CCTGGTGCCTA CGCCTGAATA AGTGTATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTGGTGTG ACATAATTGG
3241 ACAAACTACC TACAGAGATT TAAAGCTCTA AGGTAATAT AAAATTTTA AGTGTATAAT
3301 GTGTTAAACT ACTGATTCTA ATTGTTGTG TATTTAGAT TCCAACTTAT GGAACGTGATG
3361 AATGGGAGCA GTGGTGGAAAT GCCTTAAATG AGGAAAACCT GTTTGCTCA GAAGAAATGC
3421 CATCTAGTGA TGATGAGGCT ACTGCTGACT CTCAACATTC TACTCCTCCA AAAAAGAAGA
3481 GAAAGGTAGA AGACCCCAAG GACTTCCTT CAGAATTGCT AAGTTTTTG AGTCATGCTG
3541 TGTTTAGTAA TAGAACTCTT GCTTGCTTTG CTATTTACAC CACAAAGGAA AAAGCTGCAC
3601 TGCTATACAA GAAAATTATG GAAAAATATT CTGTAACCTT TATAAGTAGG CATAACAGTT
3661 ATAATCATAA CATACTGTTT TTTCTTACTC CACACAGGCA TAGAGTGTCT GCTATTAATA
3721 ACTATGCTCA AAAATTGTGT ACCTTAGCT TTTAATTG TAAAGGGTT AATAAGGAAT
3781 ATTTGATGTA TAGTGCCTTG ACTAGAGATC ATAATCAGCC ATACCACATT TGTAGAGGTT
3841 TTACTTGCTT TAAAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA AATGAATGCA
3901 ATTGTTGTG TTAACCTGTT TATTGCAGCT TATAATGGTT ACAAAATAAG CAATAGCATC
3961 ACAAAATTCA CAAATAAAGC ATTTTTCTA CTGCATTCTA GTTGTGGTTT GTCCAAACTC
4021 ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCGGGTAC

//

Figure 20: Nucleic acid sequence of the known segment of the R15/appa plasmid (including the vector sequences of pBLCAT3 (SEQ ID NO:4).

LOCUS R15/appa 6116 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa transgene with vector
 ACCESSION R15/appa
 REFERENCE 1 (bases 1 to 6116)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Rattus.
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers
 DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
 gene,
 ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appA
 reveals significant homology between pH 2.5 acid phosphatase
 and glucose-1-phosphatase
 JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)

Figure 20A:

```
MEDLINE 90368616

FEATURES          Location/Qualifiers
Source            1811..3109
                  /organism="Escherichia coli"
                  /db_xref="taxon:562"
sig_peptide       1811..1876
                  /gene="appA"
CDS               1811..3109
                  /gene="appA"
                  /standard_name="acid phosphatase/phytase"
                  /transl_table=11
                  /product="periplasmic phosphoanhydride phosphohydrolase"
                  /protein_id="AAA72086.1"
                  /db_xref="GI:145285"

mat_peptide       1877  3106
                  /gene="appA"
                  /product="periplasmic phosphoanhydride phosphohydrolase"

mutation          replace(1817.. 1819, "gcg changed to gcc")
                  /gene="appA"
                  /standard_name="A3 mutant"
                  /note="created by site directed mutagenesis"
                  /phenotype="silent mutation"
mutation          replace(3092..3094, " ccg changed to ccc")
                  /gene="appA"
                  /standard_name=" P428 mutant"
                  /note="created by site directed mutagenesis"
                  /phenotype=" silent mutation "
mutation          replace(3095..3097, " gcg changed to gct")
                  /gene="appA"
                  /standard_name=" A429 mutant"
                  /note="created by site directed mutagenesis"
                  /phenotype=" silent mutation "

DEFINITION Plasmid pBLCAT3  (bases 3109 to 6116)
ACCESSION X64409
VERSION  X64409.1  GI:58163
SOURCE   synthetic construct.
ORGANISM  synthetic construct
          artificial sequence.
REFERENCE 1  (bases 3109 to 6116)
AUTHORS  Luckow, B.H.R.
TITLE    Direct Submission
JOURNAL  Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res
          Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG
```

Figure 20B:

REFERENCE 2 (bases 3109 to 6116)
AUTHORS Luckow, B. and Schutz, G.
TITLE CAT constructions with multiple unique restriction sites
for the functional analysis of eukaryotic promoters and
regulatory elements
JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)
MEDLINE 87260024
COMMENT Promoterless CAT vector for transient transfection
experiments with eukaryotic cells. Allows the analysis of foreign
promoters and enhancers.
FEATURES Location/Qualifiers
source 3109 to 6116
/organism="synthetic construct"
/db_xref="taxon:32630"
polyA_signal 3262..3457
/note="SV40 signals"
CDS complement(4654..5514)
/codon_start=1
/transl_table=11
/gene="Amp"
/product="beta-lactamase"
/protein_id="CAA45753.1"
/db_xref="GI:58165"
BASE COUNT 1724 a 1386 c 1407 g 1599 t
ORIGIN
1 GGATCCCTT TGCTATGTAG TTTTAATGG AAATTACAAC CCATAGTGTG TTGATAAATA
61 GAGAGTCCTG TTTGGTTAA GCAACCTCTG TTTCTCATAA ACTCCATAAA AACAGGAATA
121 CTCTTGTAA CTAGCATAAC CAAAAGATT AGTGAATTGA AAACAATGTT CCCTTAGAGT
181 ATAGGTCTAA TAACCCGAA AATATTACCA TGATACTGAG CATTTGTAAG TATCTCATAG
241 CATGTAGTAT CCATAGTCCA TCAATGAGAG AGACATTAA CATGATTTTC ATTAATCAGG
301 TGGAAAAGAC ATGACAACAT TCACAGGCAC TGCACAGAAC ATAGTGGTCC ACCTTGACACA
361 TATTCACTA AACTAGGTT ATCTATTTG TTGCTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTACTTA ACCTAAGGAA CACAAAAAAAT TTTCTACATA
481 TATCCTGGTT AGAGAGTGT TAAAATAAGT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
601 TGGGAAGAAA CCATTTGGTG ACAATATTT CAAATAAAAA TAGACAAACA TAGTTAATTG
661 TAAAACATAT GTTGACCAAG CCCTTCTTT CAATAGGCTT AATGTGAATA AAATGTTAAA
721 GATTCTCTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
901 TAAGATAAAAG GTAACTGTAT ACATTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATTCTTGC CTTAAACAAC TTAAATAGAA TGATTTAAAA TATGGAGCTG
1021 TCCATGGGAC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCTT
1081 GATACTAACAA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTGCTGG GATTTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTGGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTAAGC TGTACTATTG ATCAAAGAAA TTTATTACCT TCAGTTCAA TGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAACCTTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTTA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTCTTA CACATGATT TATTCAAATT TACTCTAAC
1561 ATTGTTGAAC CATTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGGC
1621 AAAAGTCCCA GTGTGGAGTA AAGGATGCAA GATTTCCCTGC TCTGTTAAGT ATAAAATAAT

Figure 20C:

1681 AGTATGAATT CAAAGGTGCC ATTCTTCTGC TTCTAGTTAT AAAGGCAGTG CTTGCTTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTTATCT CTTCTGATTC CGTTAACCCC
1861 GCAATCTGCA TTCGCTCAGA GTGAGCCCGA CCTCAACCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCAGA
1981 CGCATGCCA ACCTGGCCGG TAAAACGTGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCGAAAAA
2101 GGGCTGCCCG CAGTCTGGTC AGGTCGCGAT TATTGCTGAT GTCGACGAGC GTACCCGTA
2161 AACAGGCAGA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCCA
2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
2281 GGATAACGCG AACGTGACTG AC CGCAGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTCGCGA ACTGGAACGG GTGCTTAATT TTCCGCAATC
2401 AAACTTGTGC CTAAACGTG AGAAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTACC
2461 ATCGGAACTC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTC TCCTGCAACA AGCACAGGGA ATGCCGGAGC CGGGGTGGGG
2581 AAGGATCACC GATTACACACC AGTGGAACAC CTTGCTAAGT TTGCATAACG CGCAATTAA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCCG CAGCCGCC ACCCCGTTAT TAGATTGAT
2701 CAAGACAGCG TTGACGCCCG ATCCACCGCA AAAACAGGGC TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATCGCCG GACACGATAC TAATCTGGCA AATCTCGGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGG TAACACGCC CGAGGTGGTG AACTGGTGT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGTTCTCCA
2941 GACTTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CGGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCACGCAT ACCCGCTTGC AGTTTGTAAAG GTATAAGGCA
3121 GTTATTGGTG CCCTTAAACG CCTGGTCTA CGCCTGAATA AGTGATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTTGGTGTG ACATAATTGG
3241 ACAAACTACC TACAGAGATT TAAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA
3301 AATGAATGCA ATTGTTGTT TTAACCTGTT TATTGCAAGT TATAATGGTT ACAAATAAAG
3361 CAATAGCATC ACAAATTTC ACAAATAAAGC ATTTTTTCA CTGCATTCTA GTTGTGGTT
3421 GTCCAAACTC ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCGGGTAC CGAGCTCGAA
3481 TTCGTAATCA TGGTCATAGC TGTTTCTGT GTGAAATTGT TATCCGCTCA CAATTCCACA
3541 CAACATACGA GCCGGAAGCA TAAAGTGTAA AGCCTGGGT GCCTAATGAG TGAGCTAACT
3601 CACATTAATT GCGTTGCGCT CACTGCCGC TTTCCAGTCG GGAAACCTGT CGTGCCAGCT
3661 GCATTAATGA ATCGGCCAAC GCGCGGGGAG AGGCGGTTG CGTATTGGC GCTCTCCGC
3721 TTCCTCGCTC ACTGACTCGC TGCGCTCGGT CGTTCGGCTG CGCGAGCGG TATCAGCTCA
3781 CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGAT AACGCAGGAA AGAACATGTG
3841 AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAGGCC GCCTGCTGG CGTTTTCCA
3901 TAGGCTCCGC CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA
3961 CCCGACAGGA CTATAAAGAT ACCAGGC GTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC
4021 TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTT CTCCCTTCGG GAAGCGTGGC
4081 GCTTCTCAA TGCTCACGCT GTAGGTATCT CAGTTGGTG TAGGTCGTT GCTCCAAGCT
4141 GGGCTGTGTG CACGAACCCC CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATATCG
4201 TCTTGAGTCC AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG
4261 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTT TTGAAGTGGT GGCCTAACTA
4321 CGGCTACACT AGAAGGACAG TATTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTCGG
4381 AAAAGAGATT GGTAGCTCTT GATCCGGCAA ACAAAACCAC GCTGGTAGCG GTGGTTTTT
4441 TGTTTGAAG CAGCAGATTA CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT
4501 TTCTACGGGG TCTGACGCTC AGTGGAACGA AAACCTACCGT TAAGGGATT TGTCATGAG
4561 ATTATCAAAA AGGATCTTC CCTAGATCCT TTTAAATTAA AAATGAAGTT TTAAATCAAT
4621 CTAAAGTATA TATGAGTAA CCTGGTCTGA CAGTTACCAA TGCTTAATCA GTGAGGCACC
4681 TATCTCAGCG ATCTGTCTAT TTCGTTCATC CATAGTTGCC TGACTCCCCG TCGTGTAGAT
4741 AACTACGATA CGGGAGGGCT TACCATCTGG CCCCAGTGT GCAATGATAC CGCGAGACCC
4801 ACGCTCACCG GCTCCAGATT TATCAGCAAT AAACCAAGCC GCGGGAGGG CCGAGCGCAG
4861 AAGTGGCTCT GCAACTTTAT CCGCCTCCAT CCAGTCTATT AATTGTTGCC GGGAAAGCTAG
4921 AGTAAGTAGT TCGCCAGTTA ATAGTTGCC CAACGTTGTT GCCATTGCTA CAGGCATCGT
4981 GGTGTCAACGC TCGTCGTTTG GTATGGCTTC ATTCAAGCTCC GGTTCCCAAC GATCAAGGCG
5041 AGTTACATGA TCCCCCATGT TGTGAAAAA AGCGGTTAGC TCCCTCGGTC CTCCGATCGT
5101 TGTCAGAAAGT AAGTTGGCCG CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATT

Figure 20D:

5161 TCTTACTGTC ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC
5221 ATTCTGAGAA TAGTGTATGC GGCAGCCGAG TTGCTCTGC CCGCGTCAA TACGGGATAA
5281 TACCGCGCCA CATAGCAGAA CTTTAAAAGT GCTCATCATT GGAAAACGTT CTTCGGGCG
5341 AAAACTCTCA AGGATCTTAC CGCTGTTGAG ATCCAGTTCG ATGTAAACCA CTCCTCCACC
5401 CAACTGATCT TCAGCATCTT TTACTTTCAC CAGCGTTCT GGGTGAGCAA AAACAGGAAG
5461 GCAAAATGCC GCAAAAAAGG GAATAAGGGC GACACGGAAA TGTTGAATAC TCATACTCTT
5521 CCTTTTCAA TATTATTGAA GCATTATCA GGGTTATTGT CTCATGAGCG GATACATATT
5581 TGAATGTATT TAGAAAAATA AACAAATAGG GGTTCCCGCG ACATTTCCCC GAAAAGTGCC
5641 ACCTGACGTC TAAGAAACCA TTATTATCAT GACATTAACC TATAAAAATA GGCGTATCAC
5701 GAGGCCCTT CGTCTCGCGC GTTTCGGTGA TGACGGTGA AACCTCTGAC ACATGCAGCT
5761 CCCGGAGACG GTCACAGCTT GTCTGTAAGC GGATGCCGGG AGCAGACAAG CCCGTCAGGG
5821 CGCGTCAGCG GGTGTTGGCG GGTGTCGGG CTGGCTTAAC TATGCAGGCAT CAGAGCAGAT
5881 TGTACTGAGA GTGCACCATA TGCGGTGTGA AATACCGCAC AGATGCGTAA GGAGAAAATA
5941 CCGCATCAGG CGCCATTGCG CATTCAAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG
6001 GGCTCTTCG CTATTACGCC AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG
6061 GGTAACGCCA GGGTTTCCC AGTCACGACG TTGTAAAACG ACGGCCAGTG CCAAGC
//

Figure 21: Nucleic acid sequence of the known segment of the R15/appa transgene used for the generation of transgenic mice (SEQ ID NO:5).

LOCUS R15/appa 3470 bp DNA SYN 15-APR-2000
 DEFINITION R15/appa transgene with vector sequences removed.
 ACCESSION R15/appa
 REFERENCE 1 (bases 1 to 3470)
 SOURCE synthetic construct.
 ORGANISM synthetic construct
 artificial sequence.
 KEYWORDS salivary proline-rich protein, acid glucose-1-phosphatase; appA
 gene; periplasmic phosphoanhydride phosphohydrolase; artificial
 sequence;
 AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
 JOURNAL Unpublished.

DEFINITION Rat salivary proline-rich protein (RP15) gene.
 ACCESSION M64793 M36414
 VERSION M64793.1 GI:206711
 SOURCE Rat (Sprague-Dawley) liver DNA.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;
 Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 1748)
 AUTHORS Lin, H.H. and Ann, D.K.
 TITLE Molecular characterization of rat multigene family
 encoding proline-rich proteins
 JOURNAL Genomics 10, 102-113 (1991)
 MEDLINE 91257817
 FEATURES Location/Qualifiers
 source 1..1748
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 /tissue_lib="cosmid genomic library"
 misc_feature 1802-1810
 /function=" consensus sequence for initiation in
 higher eukaryotes "

FEATURES Location/Qualifiers

DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
 gene,
 ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
 VERSION M58708.1 GI:145283
 SOURCE Escherichia coli DNA.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (bases 1811..3109)
 AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appA
 reveals significant homology between pH 2.5 acid phosphatase
 and glucose-1-phosphatase

Figure 21A:

JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE 90368616

FEATURES Location/Qualifiers

Source 1811..3109
/organism="Escherichia coli"
/db_xref="taxon:562"

sig_peptide 1811..1876
/gene="appA"

CDS 1811..3109
/gene="appA"
/standard_name="acid phosphatase/phytase"
/transl_table=11
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AAA72086.1"
/db_xref="GI:145285"

mat_peptide 1877 3106
/gene="appA"
/product="periplasmic phosphoanhydride phosphohydrolase"

mutation replace(1817.. 1819, "gcg changed to gcc")
/gene="appA"
/standard_name="A3 mutant"
/note="created by site directed mutagenesis"
/phenotype="silent mutation"

mutation replace(3092..3094, " ccg changed to ccc")
/gene="appA"
/standard_name=" P428 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "

mutation replace(3095..3097, " gcg changed to gct")
/gene="appA"
/standard_name=" A429 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "

polyA_signal 3262..3457
/note="SV40 signals"

BASE COUNT 1065 a 721 c 735 g 949 t
ORIGIN

1	GGATCCCTT	TGCTATGTAG	TTTTAATGG	AAATTACAAC	CCATAGTGTG	TTGATAAATA
61	GAGAGTCCTG	TTGGTTAA	GCAACCTCTG	TTTCTCATAA	ACTCCATAAA	AACAGGAATA
121	CTCTTGTTT	CTAGCATAAC	CAAAAGATT	AGTGAATTGA	AAACAATGTT	CCCTTAGAGT
181	ATAGGTCTAA	TAACCCCGAA	AATATTACCA	TGATACTGAG	CATTTGTAAG	TATCTCATAG
241	CATGTAGTAT	CCATAGTCCA	TCAATGAGAG	AGACATTTAA	CATGATTTTC	ATTAATCAGG

Figure 21B:

301 TGGAAAAGAC ATGACAAACAT TCACAGGCAC TGACAGAAC ATAGTGGTCC ACCTTGCACA
361 TATTCACTA AACTAGGTTT ATCTATTTG TTGCTTCTC TAACATCTCT GCAATGAAGC
421 AGGTCAACAG TGCCACATAT CCTTTACTTA ACCTAAGGAA CACAAAAAAT TTTCTACATA
481 TATCCTGGTT AGAGAGTGCT TAAATAACT TTTCCAAGAA TGGAAAAGAA ATGTTCTGAC
541 TTAACAATTA AGACAGTATT TATTTAAAGC AAGAAATATG AGGCACACAA GAAAATATTT
601 TGGGAAGAAA CCATTTGGTG ACAATATTT CAAATAAAA TAGACAAACA TAGTTAATTG
661 TAAAACATAT GTTGACCAAG CCCTCTTT CAATAGGCTT AATGTGAATA AAATGTAAA
721 GATTCTCTT GGGTGGCTGC AAATTGTCCA CGAATAAGAC AAAATATAAA AATAAGGACT
781 GAGTCTCACA AAATGAAAAG GAAATATATT CAGAAAGAGA ATCTTGAGAG AATGTGTTGT
841 CACAAATTAA AGAAAACCTG TGGTGAATGA CATCCTGAGG CCTGAGCTAT TACTGACATT
901 TAAGATAAAG GTAACTGTAT ACATTGTCC CATTGAGGGG ACAAGAAAGC TGCTCTCATG
961 TTCAGCTCTA TAATTCTTGC CTTAAACAAC TTAAATAGAA TGATTTAAAAT TATGGAGCTG
1021 TCCATGGACC TTTGAAATAT AAAATAGTCA AGCAACTTAT CAAGGAATTA CAGATTCCCT
1081 GATACTAACCA CAGGTAAATC CCACACGTGT TTTGAGACTA CATTGCTGTT GATTTATTG
1141 ATGTAATAGG TCACATGTTT TTCGGGCCAA TGTTGCTGTT ATTGGTTAC TTCAAGAGAA
1201 TAGTGGCAAC TGATGCTATG TATTCTAGGG GTTTGAAGTG ATGTTTCATG ATTGAAATTT
1261 GTAAAAGAAT AACATCATCA TTCTTAACAA TAGAACATAT AAAGTCACAC AGAAGTGACA
1321 GTGTTAAGC TGACTATTG ATCAAAGAAA TTTATTACCT TCAGTTCAA TGAAATAAT
1381 TACTGATAAT ACAAACATGT GTGAACACAC ACTAACCTTA TCCAAATGCA CAGTGATACA
1441 CAGAAAATAT TAGCAAGTAG AATGCAATAT TTATATAACG ATTGTATTAA TCAATCAATT
1501 GTATGTATCA ATATATGGGC TATTTCCTTA CACATGATT TATTCAAATT TACTCTAAC
1561 ATTGTTAAC CATTAGAAA AGGCATACTG GCAACTTTTC CTTACCTCAT CCAGCTGGC
1621 AAAAGTCCC GTGTGGAGTA AAGGATGCAA GATTTCCTGC TCTGTTAAGT ATAAAATAAT
1681 AGTATGAATT CAAAGGTGCC ATTCTCTGC TTCTAGTTA AAAGGCAGTG CTTGCTCTT
1741 CCAGCACAGA TCTGGATCTC GAGGAGCTTG GCGAGATTT CAGGAGCTAA GGAAGCTAAA
1801 AGCCGCCACC ATGAAAGCCA TCTTAATCCC ATTTTATCT CTTCTGATTG CGTTAACCCC
1861 GCAATCTGCA TTGCTCAGA GTGAGCCGGA GCTGAAGCTG GAAAGTGTGG TGATTGTCAG
1921 TCGTCATGGT GTGCGTGCTC CAACCAAGGC CACGCAACTG ATGCAGGATG TCACCCAGA
1981 CGCATGGCCA ACCTGGCCGG TAAAACCTGGG TTGGCTGACA CCGCGCGGTG GTGAGCTAAT
2041 CGCCTATCTC GGACATTACC AACGCCAGCG TCTGGTAGCC GACGGATTGC TGGCAAAAAA
2101 GGGCTGCCCG CAGTCTGGTC AGGTGCGAT TATTGCTGAT GTGACGAGC GTACCCGTA
2161 AACAGGCAGA GCCTTCGCCG CCGGGCTGGC ACCTGACTGT GCAATAACCG TACATACCA
2221 GGCAGATACG TCCAGTCCCG ATCCGTTATT TAATCCTCTA AAAACTGGCG TTTGCCAACT
2281 GGATAACCGC AACGTGACTG ACGCGATCCT CAGCAGGGCA GGAGGGTCAA TTGCTGACTT
2341 TACCGGGCAT CGGCAAACGG CGTTTGCAGA ACTGGAACCG GTGCTTAATT TTCCGCAATC
2401 AAACTTGTGC CTAAACGTG AGAAACAGGA CGAAAGCTGT TCATTAACGC AGGCATTAC
2461 ATCGGAACTC AAGGTGAGCG CCGACAATGT CTCATTAACC GGTGCGGTAA GCCTCGCATC
2521 AATGCTGACG GAGATATTTC TCCTGCAACA AGCACAGGGAA ATGCCGGAGC CGGGGTGGGG
2581 AAGGATCACC GATTCAACACC AGTGGAACAC CTTGCTAAGT TTGCATAACG CGCAATTAA
2641 TTTGCTACAA CGCACGCCAG AGGTTGCCG CAGCCGCC ACCCCGTTAT TAGATTGAT
2701 CAAGACAGCG TTGACGCCCG ATCCACCGCA AAAACAGGCG TATGGTGTGA CATTACCCAC
2761 TTCAGTGCTG TTTATGCCG GACAGGATAC TAATCTGGCA AATCTCGCG GCGCACTGGA
2821 GCTCAACTGG ACGCTTCCCG GTCAGCCGGA TAACACGCCG CCAGGTGGTG AACTGGTGT
2881 TGAACGCTGG CGTCGGCTAA GCGATAACAG CCAGTGGATT CAGGTTTCGC TGGTCTCCA
2941 GACTTACAG CAGATGCGTG ATAAAACGCC GCTGTCATTA AATACGCCGC CGGGAGAGGT
3001 GAAACTGACC CTGGCAGGAT GTGAAGAGCG AAATGCGCAG GGCATGTGTT CGTTGGCAGG
3061 TTTTACGCAA ATCGTGAATG AAGCACCGAT ACCCGCTTGC AGTTGTAAG GTATAAGGCA
3121 GTTATTGGTG CCCTAAACG CCTGGTGCTA CGCCTGAATA AGTATAATA AGCGGATGAA
3181 TGGCAGAAAT TCGCCGGATC TTTGTGAAGG AACCTTACTT CTGTTGTTG ACATAATTGG
3241 ACAAACCTACC TACAGAGATT TAAAAAACCT CCCACACCTC CCCCTGAACC TGAAACATAA
3301 AATGAATGCA ATTGTTGTT TTAACCTGTT TATTGCAAGCT TATAATGGTT ACAAAATAAG
3361 CAATAGCATC ACAAAATTCA CAAATAAAAGC ATTTTTTCA CTGCATTCTA GTTGTGGTTT
3421 GTCCAAACTC ATCAATGTAT CTTATCATGT CTGGATCGAT CCCCGGGTAC

//

Figure 22: Nucleic acid sequence of the SV40/APPA+intron plasmid (SEQ ID NO:6).

LOCUS SV40/APPA 5421 bp DNA CIRCULAR SYN 14-APR-2000
DEFINITION Ligation of SV40 promoter/enhancer into CAT/APPA+intron
ACCESSION SV40/APPA
REFERENCE 1 (bases 1 to 5421)
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
KEYWORDS SV40 promoter/enhancer, acid glucose-1-phosphatase; appA gene;
periplasmic phosphoanhydride phosphohydrolase; artificial
sequence;
AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
JOURNAL Unpublished.
DEFINITION E. coli periplasmic phosphoanhydride phosphohydrolase (appA)
gene,
ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375
VERSION M58708.1 GI:145283
SOURCE Escherichia coli DNA.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (bases 40 1337)
AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
TITLE The complete nucleotide sequence of the Escherichia coli gene appA
reveals significant homology between pH 2.5 acid phosphatase
and glucose-1-phosphatase
JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
MEDLINE 90368616
FEATURES Location/Qualifiers
Source 40 1337
/organism="Escherichia coli"
/db_xref="taxon:562"
sig_peptide 40..105
/gene="appA"
CDS 40 1337
/gene="appA"
/standard_name="acid phosphatase/phytase"
/transl_table=11
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AAA72086.1"
/db_xref="GI:145285"
mat_peptide 106 1334
/gene="appA"

Figure 22A:

```
/product="periplasmic phosphoanhydride phosphohydrolase"

mutation    replace(46.. 48, "gcg changed to gcc")
/gene="appA"
/standard_name="A3 mutant"
/note="created by site directed mutagenesis"
/phenotype="silent mutation"
mutation    replace(1320..1322, " ccg changed to ccc")
/gene="appA"
/standard_name=" P428 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "
mutation    replace(1323..1325, " gcg changed to gct")
/gene="appA"
/standard_name=" A429 mutant"
/note="created by site directed mutagenesis"
/phenotype=" silent mutation "

DEFINITION Plasmid pBLCAT3 (bases 2200 to 4924)
ACCESSION X64409
VERSION X64409.1 GI:58163
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequence.
REFERENCE 1 (bases 2200 to 4924)
AUTHORS Luckow,B.H.R.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1992) B.H.R. Luckow, German Cancer Res
Center, Im Neuenheimer Feld 280, W-6900 Heidelberg, FRG
REFERENCE 2 (bases 2200 to 4924)
AUTHORS Luckow,B. and Schutz,G.
TITLE CAT constructions with multiple unique restriction sites
for
regulatory
elements
JOURNAL Nucleic Acids Res. 15 (13), 5490 (1987)
MEDLINE 87260024
COMMENT Promoterless CAT vector for transient transfection
experiments
with eukaryotic cells. Allows the analysis of foreign
promoters and enhancers.
FEATURES Location/Qualifiers
source 2200 to 4924
/organism="synthetic construct"
/db_xref="taxon:32630"

SV40 t intron 1380..1993
/note="SV40 signals"
polyA_signal 1990..2230
/note="SV40 signals"
CDS complement(3471..4317)
/codon_start=1
/transl_table=11
/gene="Amp"
/product="beta-lactamase"
/protein_id="CAA45753.1"
/db_xref="GI:58165"
```

Figure 22B:

SV40 promoter/enhancer 5023..5402
/note="SV40 signals"

BASE COUNT 1413 a 1321 c 1331 g 1355 t
ORIGIN

1 CGAGATTTTC AGGAGCTAAG GAAGCTAAAA GCCGCCACCA TGAAAGCCAT CTTAATCCCC
61 TTTTTATCTC TTCTGATTCC GTTAACCCCG CAATCTGCAT TCGCTCAGAG TGAGCCGGAG
121 CTGAAGCTGG AAAGTGTGGT GATTGTCAGT CGTCATGGTG TGCGTGCCTC AACCAAGGCC
181 ACGCAACTGA TGCAGGATGT CACCCAGAC GCATGGCCAA CCTGGCCGGT AAAACTGGGT
241 TGGCTGACAC CGCGNGGTGG TGAGCTAATC GCCTATCTCG GACATTACCA ACGCCAGCGT
301 CTGGTAGCCG ACGGATTGCT GGCAGAAAAG GGCTGCCCG AGCTCTGGTCA GGTCGCGATT
361 ATTGCTGATG TCGACGAGCG TACCCGTAAA ACAGGCGAAG CCTTCGCCGC CGGGCTGGCA
421 CCTGACTGTG CAATAACCAGT ACATACCCAG GCAGATACTG CCAGTCCCAG TCCGTTATTT
481 AATCCTCTAA AAACGGCGT TTGCCAACTG GATAACGCGA ACGTGACTGA CGCGATCCTC
541 AGCAGGGCAG GAGGGTCAAT TGCTGACTTT ACCGGGCATC GGCAAACGGC GTTTCGCGAA
601 CTGGAACGGG TGCTTAATT TCCGCAATCA AACTTGTGCC TTAAACGTGA GAAACAGGAC
661 GAAAGCTGTT CATTAAACGCA GGCATTACCA TCGGAACTCA AGGTGAGCGC CGACAATGTC
721 TCATTAACCG GTGCGGTAAG CCTCGCATCA ATGCTGACGG AGATATTCT CCTGCAACAA
781 GCACAGGGAA TGGCGGAGCC GGGGTGGGAGGAGATCACCG ATTACACCCA GTGGAACACC
841 TTGCTAAGTT TGCATAACGC GCAATTCTTAT TTGCTACAAC GCACGCCAGA GTTTCGCCGC
901 AGCCGCCCA CCCCGTTATT AGATTGATC AAGACAGCGT TGACGCCCA CCACCGCAAA
961 AACAGGCAGTA TGGTGTGACA TTACCCACTT CAGTGTGTT TATGCCCGA CACGATACTA
1021 ATCTGGAAA TCTCGGGCGC GCACTGGAGC TCAACTGGAC GCTTCCCAGT CAGCCGGATA
1081 ACACGCCGCC AGGTGGTGAAT CTGGTGTGTT AACGCTGGCG TCGGCTAAGC GATAACAGCC
1141 AGTGGATTCA GTTTCGCTG GTCTTCCAGA CTTTACAGCA GATGCGTGT AAAACGCCGC
1201 TGTCTTAAAT TACGCCGCC GGAGAGGTGA AACTGACCCCT GGCAAGGATGT GAAGAGCGAA
1261 ATGCGCAGGG CATGTGTTCG TTGGCAGGTT TTACGCAAAT CGTGAATGAA GCACGCATAC
1321 CCGCTTGCAG TTTGTAAGGC AGTTATTGGT GCCCTTAAAC GCCTGGTGCT ACGCCTGAAT
1381 AAGTGATAAT AAGCGGATGA ATGGCAGAAA TTCGCCGGAT CTTGTGAAG GAACCTTACT
1441 TCTGTGGTGT GACATAATTG GACAAACTAC CTACAGAGAT TTAAAGCTCT AAGGTAATA
1501 TAAAATTCTT AAGTGTATAA TGTGTAAAC TACTGATTCT ATTGTTGT GTATTTTAGA
1561 TTCCAACCTA TGGAACTGAT GAATGGGAGC AGTGGTGGAA TGCCTTAAAT GAGGAAAACC
1621 TGTTTGCTC AGAAGAAATG CCATCTAGTG ATGATGAGGC TACTGCTGAC TCTAACATT
1681 CTACTCCTCC AAAAAAGAAG AGAAAGGTAG AAGACCCCA GGACTTTCT TCAGAATTGC
1741 TAAGTTTTT GAGTCATGCT GTGTTAGTA ATAGAACTCT TGCTTGCTTT GCTATTAC
1801 CCACAAAGGA AAAAGCTGCA CTGCTATACA AGAAAATTAT GGAAAATAT TCTGTAACCT
1861 TTATAAGTAG GCATAACAGT TATAATCATA ACATACTGTT TTTCTTACT CCACACAGGC
1921 ATAGAGTGTC TGCTATTAAT AACTATGCTC AAAAATTGTG TACCTTCTAGC TTTTTAATT
1981 GTAAAGGGGT TAATAAGGAA TATTGATGT ATAGTGCCTT GACTAGAGAT CATAATCAGG
2041 CATACCACAT TTGTTAGAGGT TTTACTTGCT TTAAAAAAACC TCCCACACCT CCCCCTGAAC
2101 CTGAAACATA AAATGAATGC AATTGTTGTT GTTAACTTGT TTATTGCAAG TTATAATGGT
2161 TACAAATAAA GCAATAGCAT CACAAATTTC ACAAAATAAG CATTTCCTC ACTGCATTCT
2221 AGTTGTGGTT TGCCAAACT CATCAATGTA TCTTATCATG TCTGGATCGA TCCCCGGGTA
2281 CCGAGCTCGA ATTGTAATC ATGGTCATAG CTGTTTCTG TGTGAAATTG TTATCCGCTC
2341 ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTAA AGCCTGGGG TGCCTAATGA
2401 GTGAGCTAAC TCACATTAAT TGCGTGCAGC TCACTGCCCG CTTTCCAGTC GGGAAACCTG
2461 TCGTGCAGC TGCTTAATG AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG
2521 CGCTCTTCCG CTTCTCGCT CACTGACTCG CTGCGCTCG TCGTTGGCT GCGCGAGCG
2581 GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG AATCAGGGGA TAACGCAGGA
2641 AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC GTAAAAAGGC CGCGTTGCTG
2701 GCGTTTTCC ATAGGCTCCG CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG
2761 AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAAGCGT TTCCCCCTGG AAGCTCCCTC
2821 GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT TCTCCCTTCG
2881 GGAAGCGTGG CGCTTCTCA ATGCTCACGC TGTAGGTATC TCAGTTCGGT GTAGGTCGTT
2941 CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTCAAGC CGACCGCTG CGCCTTATCC
3001 GGTAACTATC GTCTTGAGTC CAACCCGGTA AGACACGACT TATGCCACT GGCAGCAGCC
3061 ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG

Figure 22C:

3121 TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT GCTGAAGCCA
3181 GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA AACAAACCAC CGCTGGTAGC
3241 GGTGGTTTTT TTGTTGCAA GCAGCAGATT ACGCGCAGAA AAAAGGATC TCAAGAAGAT
3301 CCTTTGATCT TTTCTACGGG GTCTGACGCT CACTCCAACC AAAACTCACG TTAAGGGATT
3361 TTGGTCATGA GATTATCAA AAGGATCTTC ACCTAGATCC TTTAAATTA AAAATGAAGT
3421 TTTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG ACAGTTACCA ATGCTTAATC
3481 AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTCAT CCATAGTTGC CTGACTCCCC
3541 GTCGTGAGA TAATCACGAT ACGGGAGGGC TTACCATCTG GCCCCAGTGC TGCAATGATA
3601 CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG
3661 GCGGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC
3721 CGGGAAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTGC GCAACGTTGT TGCCATTGCT
3781 ACAGGCATCG TGGTGTACCG CTCGTCGTTT GGTATGGCTT CATTCAAGCTC CGGTTCCCAA
3841 CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT
3901 CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTAT CACTCATGGT TATGGCAGCA
3961 CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC
4021 TCAACCAAGT CATTCTGAGA ATAGTGTATG CGGCGACCGA GTTGCTCTTG CCCGGCGTCA
4081 ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT
4141 TCTTCGGGGC GAAAACCTCTC AAGGATCTT CCGCTGTGA GATCCAGTTC GATGTAACCC
4201 ACTCGTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA
4261 AAAACAGGAA GGCAAAATGC CGCAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA
4321 CTCATACTCT TCCTTTTCA ATATTATTGA AGCATTATC AGGTTATTG TCTCATGAGC
4381 GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTCCC
4441 CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA TGACATTAAC CTATAAAAT
4501 AGGCGTATCA CGAGGCCCTT TCGTCTCGCG CGTTTCGGTG ATGACGGTGA AAACCTCTGA
4561 CACATGCAGC TCCCAGGAGAC GGTACACAGCT TGTCTGTAAAG CGGATGCCGG GAGCAGACAA
4621 GCCCCGTCAAG GCGCGTCAGC GGGTGTGGC GGGTGTGGG GCTGGCTTAA CTATGCGCA
4681 TCAGAGCAGA TTGTACTGAG AGTGCACCAT ATGCGGTGTG AAATACCGCA CAGATGCGTA
4741 AGGAGAAAAT ACCGCATCAG GCGCCATTG CCATTCAAGGC TCGCACAATG TTGGGAAGGG
4801 CGATCGGTGC GGGCCTCTTC GCTATTACGC CAGCTGGCGA AAGGGGGATG TGCTGCAAGG
4861 CGATTAAGTT GGGTAACGCCC AGGGTTTCC CAGTCACGAC GTTGTAAAAC GACGGCCAGT
4921 GCCAAGCTTT ACACTTTATG CTTCCGGCTC GTATGTTGTG TGGAAATTGTG AGCGGATAAC
4981 AATTCACAC AGGAAACAGC TATGACCATG ATTACGAATT CGGCGCAGCA CCATGGCCTG
5041 AAATAACCTC TGAAAGAGGA ACTTGGTTAG GTACCTCTG AGGCGGAAAG AACCAAGCTGT
5101 GGAATGTGTG TCAGTTAGGG TGTGGAAAGT CCCCAGGCTC CCCAGCAGGC AGAAGTATGC
5161 AAAGCATGCA TCTCAATTAG TCAGCAACCA GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG
5221 GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CATAGTCCCCG CCCCTAACTC
5281 CGCCCATCCC GCCCCTAACT CCGCCCAAGTT CCGCCCCATTG TCCGCCCCAT GGCTGACTAA
5341 TTTTTTTAT TTATGCAGAG GCCGAGGCCG CCTCGGCCTC TGAGCTATTG CAGAAGTAGT
5401 GAGGAGGCTC GAGGAGCTTG G G

//

Figure 23. The nucleic acid sequence of the Lama2/APPA transgene used for the generation of transgenic mice and transgenic pigs (SEQ ID NO: 7)

LOCUS transgene 17732 bp DNA SYN 14-APR-2000
DEFINITION Lama-appA cut XhoI..20623 to NotI..17732
ACCESSION transgene
KEYWORDS parotid secretory protein; acid glucose-1-phosphatase; appA gene; periplasmic phosphoanhydride phosphohydrolase; artificial sequence; cloning vector
REFERENCE 1 (bases 1 to 17732)
AUTHORS Golovan, S., Forsberg, C.W., Phillips, J.
JOURNAL Unpublished.

FEATURES
DEFINITION *M. musculus* Psp gene for parotid secretory protein.
ACCESSION X68699
VERSION X68699.1 GI:53809
SOURCE house mouse.
ORGANISM *Mus musculus*
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; *Mus*.
REFERENCE 1 (bases 3777 to 5332;)
AUTHORS Svendsen, P., Laursen, J., Krogh-Pedersen, H. and Hjorth, J.P.
TITLE Novel salivary gland specific binding elements located in the PSP proximal enhancer core
JOURNAL Nucleic Acids Res. 26 (11), 2761-2770 (1998)
MEDLINE 98256451
REFERENCE 2 (bases 7147 to 12653; 13952 to 17731)
AUTHORS Mikkelsen, T.R.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1992) T.R. Mikkelsen, Department of Molecular Biology, University of Aarhus, CF Møllers Alle 130, 8000 Aarhus, DENMARK
REFERENCE 3 (bases 7147 to 12653; 13952 to 17731)
AUTHORS Laursen J, Hjorth JP
TITLE A cassette for high-level expression in the mouse salivary glands.
JOURNAL Gene 1997 Oct 1;198(1-2):367-72
MEDLINE 9370303

FEATURES Location/Qualifiers
source 1..to 12653; 13952 to 17731
/organism="Mus musculus"
/strain="C3H/As"
/db_xref="taxon:10090"
/chromosome="2"
/map="Estimate: 69 cM from centromere"
/clone="Lambda YP1, Lambda YP3, Lambda YP7"
/clone_lib="Lambda-PHAGE (Lambda L47.1)"
/germline
/note="Allele: b"

misc_feature 3777-5332
/gene="PSP"
/function="salivary gland specific positive acting regulatory region"
enhancer 7147..8724

Figure 23A:

```
        /evidence=experimental
exon      11778..11824
        /gene="Psp"
        /note="exon a"
        /number=1
        /evidence=experimental
exon      12626.. 14190
        /gene="Psp"
        /note="exon b fused with exons h and i"
misc_feature 12644-12652
        /function=" consensus sequence for initiation in higher
eukaryotes "
misc_feature 13952-13965
        /function=" M13mp18 polylinker"
```

DEFINITION *E. coli* periplasmic phosphoanhydride phosphohydrolase (appA) gene,

ACCESSION M58708 L03370 L03371 L03372 L03373 L03374 L03375

VERSION M58708.1 GI:145283

SOURCE Escherichia coli DNA.

ORGANISM Escherichia coli

Bacteria; Proteobacteria; gamma subdivision;

Enterobacteriaceae;

Escherichia.

REFERENCE 1 (bases 12653..13951)

AUTHORS Dassa,J., Marck,C. and Boquet,P.L.

TITLE The complete nucleotide sequence of the *Escherichia coli* gene appA reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase

JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)

MEDLINE 90368616

FEATURES Location/Qualifiers

Source 12653..13951
/organism="Escherichia coli"
/db_xref="taxon:562"

sig_peptide 12653..12718

/gene="appA"

CDS 12653 13951
/gene="appA"
/standard_name="acid phosphatase/phytase"
/transl_table=11
/product="periplasmic phosphoanhydride
phosphohydrolase"
/protein_id="AAA72086.1"
/db_xref="GI:145285"

Figure 23B:

```
mat_peptide      12719 13948
/gene="appA"
/product="periplasmic phosphoanhydride
phosphohydrolase"

mutation      replace(12659.. 12661,"gcg changed to gcc")
/gene="appA"
/standard_name="A3 mutant"
/note="created by site directed mutagenesis"
/citation=[3]
/phenotype="silent mutation"
mutation      replace(13934..13936," ccg changed to ccc")
/gene="appA"
/standard_name=" P428 mutant"
/note="created by site directed mutagenesis"
/citation=[3]
/phenotype=" silent mutation "
mutation      replace(13937..13939," gcg changed to gct")
/gene="appA"
/standard_name=" A429 mutant"
/note="created by site directed mutagenesis"
/citation=[3]
/phenotype=" silent mutation "
```

BASE COUNT 4719 a 4125 c 4168 g 4719 t

ORIGIN

1 TCGAGAGTAT CTTTGTCAAGC TGTGCCTCCA ACAAAAGGGGT ACTGTTGCC ACATAGAAAG
61 ATCTAAACTA ATTAATTAAAT CCCTCACCCG CAAATCTTTC AGTCACTAAG TTAGCACGAT
121 TGTTGAACAA GTTCTCCAAA GGAGAGATAC AGATGAGTGC GTATAGGGTG GACCTGGCTG
181 CTGAGGAGAC ACCTGCATCT GACTAAGAAG AGCCACGGTG TTAGTTGAAT GGTGTGGAGT
241 AGGGTGGTTC TGTGGGACAG TAGAAAATCG AGAGGCATGT GCCGTTTAGT GAACTGATGG
301 AAGCTACCCC AAACGACAGA GATTGTCAGT CAGGCCAATC CGTTTCGAGT TTGATGGCA
361 GCCGGACAGT GAGACAGACA CACCTACTCA GTTGGAGGAA GGATGAGAAC AATGCCAGC
421 AGGGATTGAG AGACCCCTGAC AGGCCAAGG CCCTAACACA CACACCTACC ACCTCACTTG
481 ACAAAAGCTGC CAAAGACCAA AGACTTGTTC TCCATTAGAA ATGACAGCTG GCTTGGACCCG
541 ACAGCATAAT AAGCAGAGTG TACTCTGATT GGAGAACTTT AAIGTGTTC ATTCACTATT
601 ATAAAAGGAC AGTATTACAG ATTTTGTGT ACACTGCTGT TACATGTGGG GCAGTGTGTC
661 TTTAAGTAGG GTAAAGTACT CTTTAAAAAT GGGTCCTAGA TATTTTTTCC TTTAACTCAA
721 GTCTCTTACT GTTTAAATGA TTTTTATTTT GTTTAATATG GAGGAAAAAG AAGCGTAAAT
781 GGACAATATA TATTTAGAGA AAGATGGTA GCTGTCAGAA AAATATGCAA ATCAAAATCA
841 CACCAAGACT GCAGCACACC CCTGTCAGAT GGCTGTGATC AAGAAAATAA ATGACAATGA
901 GTGGTGGTGA AGATGTACTA AAGGGAAACA CACACACACA CACACACACA CACACACACA
961 CACACTGGAG CAACCACTGT GGAAATCAGT ATGAATGGTC CTCAAAACC TGAAGATAGA
1021 GCGGGGCGTG GTGGCATACA CTTTTATTCC CAGCACTGGG GAGGCAGAGG CAGGTGGATC
1081 TCTGAGTTCC AGGCCAGCCT GGTCTATAGC ACAGGTTCTA GGACAGCCAG GGCTACACAG
1141 AAAAACCTG CTTGATTAA ACCAAACCAA ACCAAACCAA ACCAAACCAA ACCAAACCAA
1201 ACCAAACCAA ACCAAACCAAG ACCAAACCAA AACACTGAAG ATAGAACTTC AGTATTCCAT
1261 TCCTAGATAT ATACCCAAATG GAGACTAAGT CAGCAAGACA CCTGCACAGC CATGTTCACT
1321 ACTACACTGT TCACCCACAGC CAGGCTGTGG AACCCAGCCTG AGTGTCCATG ATAAATGAAT

Figure 23C:

1381 GGATAGGTAA CTTCAAGGT AAATGGACTC TGCTGTGTAC ATGCCTCACA TTCTGTTTAT
1441 TCATTTTCCT TTATGAGGTG TCCATTCAAG AGTCACATGG TAGTTCTATT TTCAGTCTC
1501 TGAAGATACT AACTGGTCC CCACAGTTA CACTTTATC AGCAGTGAAT AAGGGTTCC
1561 CTATCCTTAC CATCATTGT TGTAATTCTT CTTGATGACC CTCTTCTGA CAGGGATAGG
1621 ATGTAATATC AGTGTGAGGA AGTACAACTT GTTTCTAAG TATTTATTGG CCCCTTGCA
1681 TTCTTCTTT GAAAATGTC GGTTCCCTGAC ATCTGCTCAG GTATTCAATTG GATGTTGTT
1741 CTTGGTGTG TGAGTTCTTA TGAATTCTAG ATGTTAAATC CCTGCCTGTG GTTCTCTCCC
1801 ATTCTGTAGG CTGCCCTCTC ACCCTGGCAA TTGTTGTCCT TGTTTGCAG AAACTTTGA
1861 CTTCATGGAA TCTCATTGT CAGTTTCCC TCCTCTGCTA TAGCCTGAGC TAATGCACTG
1921 GTTTTACAG AGCCCTGGTC TATGCCTTTA TCCTCCTCTG GCAGCTTCGG AGTTTCATT
1981 CTTACATTTA GATCTTGAT CCACTTGAA CAAGTTTGG AGCAGGGTGA GAGATACGAA
2041 TCTAGTTCCA TTCTTCCATA TGTGATCCTA GTTACATAG CATGTTGGT TGAAGAGGTT
2101 TTATTTTATT TTTAAATAAT GTGTCAAAA AAACGAGGTG GTTGTAGCAG TGTGGATTG
2161 TTTCTTGTG CTTGATCTA CAGGTCTTGT TTGTTGTCAG TCTCATGATG TTTTATGCT
2221 ATGGCTCTGT CATACTGCT GAGGTCAAGT ATTGTGATAT ACCTTCAGTA TTGCTCCCTC
2281 AGACTCAGGT TTGCTTGGC CAGGAGTCAT CTTACTCAGT GCTCTTAGAG CTCCCCCAGC
2341 ATGTAGCTGC TACTATTCTT AGTTGATAAA TCAGGAAACT GGGGCTCAGA GAGATTAAC
2401 GTCTTGAAC ACTTCTGGGG AGGTGAAACG TGGAGACACT AAACGTGTT TACCTGTAC
2461 TGCTCCAGTA GCTGTCGGGT GCTGGGCTAC AGCAAAGCAC CTATACTATA TATTACTCAG
2521 GAGGTGGAAA AACTCAGCCT CCCTTGGGT TCCCAAGCTC CCAGGGTGTCC AGTCACTGCT
2581 GGAAACCTCA TGGAGTCTGA AAGGAAGGGT TGAGGGTACA TGGGGCAGCG ATGAGGAGCC
2641 TGGGGCTGGG ATCTCCAAA CACCTGGATA TCCAGATGCC ACTGGGTCAG GGGGAGTTGG
2701 GAACAGAGTT GGGATGTCCA TGGACCTGTG ACAAGGCCAG GGCCAGGGGG AGGATAACTC
2761 TGGCTTACT AATTGCGAA AGTCCTTAGC TTAGCAGCAG TTGTTCTGGGA GCACAGAGGG
2821 GCCTCTGTA AGAGGCTCAG GCAGTGCCGC TCTGTAGGCG AAGGTCTTCT CCATGTTCCC
2881 CATGGTGGTT CTTGATGAAA GAGACAGTCC TTGGCTCCAA ACTGTTTAT TGATTGTC
2941 TTGTGGAAA TGGGTGCACA CCACCTCTC AGGGTGGACC AGAGATCAA TACCTTTGC
3001 AGGGAGGAAT ATCTGGGAAG GGACGCTTAC TGGCTAAACC CTCAGGGCCT CTAGATACAT
3061 CATTAGCATG GAGAACTCTG TTCTGGCTA CATGACCACA GGCCACATT CCACAAGCCA
3121 CATGTGGAA GTGTGGCACA TGTTCTAGGC CAGGAATCTG TAGGGAGCG TGGAGCCACC
3181 TACCATCCA GGTGGGTGCC TGGGTGCCAG GGACCTGAA CCCGCTCAAC CTTACCAAGT
3241 TTCTGGCAG GGTCCACTGT CCTACACAGA AGCTGGAGGA GGTGTGAGGG TTGTTCTTT
3301 GTGGAATGTC CCATGCTGCT TGGGGCTCAG TTTCTCCACC TGTACCTCAT TGGTTGGGT
3361 ATAAAAAGTG GGGATACTTT ATTATTCTCT GACTCGGTCC TGAGGAAAAA GCATCGTGGC
3421 AGTCCAGGAA CCACACCCCTG AGGTTCTGC ACTGAAGGGG CTCCTTAAGT CTCTGGAGTC
3481 TCTCCCTTC ACAGAGCTGC CAAAGTCTAG GTTCTTTGA GGATAACAGA GCCATGTTG
3541 GTAAGCAGAC AACAGCATTT GTTAACTCAA CCTCTTTTG TCAGCTCCCT TTCAAAAC
3601 AAGTTGAGAC ACCATGCTGG CTTGAGGAAG ACTTCTAAAG CCAGACAAC GTGCAAGGAA
3661 GAAGAAGAAG GGGCAAGTGG AGTTAGCCTG GATGTAGGCC TCAAAGTCTC CAGAGACCAG
3721 CCATGAAGGC TCAAGTGGAG GGCAAGACCT GCAGCAGCCA AGCATCTGGC AGGAGAGGAT
3781 CCTGGGAACC CCTCTACCAT GACACACATT CTTCTGAG GTCACACTTA ATAGGCCATT
3841 TCTTATTG TGTTCTGTG CGAGATTAAT GAGGTGTTAT GCTGGAACA
3901 GAAAGTTATA TAAAAACAAG TCCCCCCCCC TTGTCCTGC TGCTAAGAAT GTAGCAGAAA
3961 TTGTCCTCAAG TGTCCTCTA ATCAGAAACA ATAAAGGTCT CCTTGGATTC AAGCCCTCCA
4021 GTTCTCTCCT TCCCTGCTGA GCCTTGGACA CCCATACAA CCTCTGGAT GCTACAGCTC
4081 TGGGCAGAGA CTCCAAGGTG GGGAGAGACT GATGGTACAA AAGAAAATA CTTGTTGGG
4141 GGTACACCCA CCTCTCTGCTC TGTGTGGTTC CTGCAGTCAG TCCTGCAGAC AGGCCCTCAG
4201 TGGGTCTTCC ATGGGAACA CGCAGAGGGA GGCAATGGAT GGGAAATACCC ACACCCGGT
4261 TAGTTTACCC CGGGCATGCT CTCTGCTCTT CATCCCTCT CTCGCCTCTG CCACGGCTTT
4321 CTCTGCAGGA ATCATATCTT CATATTGGCC CACAGGTGTT CTCCTCACCC TAGCTATGAT
4381 GTTTACTTTA GAGTGACCTT AGCAGGGCTG GTGGGAATGA GTTCTAGAAG GCTCACGGAG
4441 ATGCTAGGGA AGAAACGTCT TCTAACTACT GAGGTACTA AGTTCCTGGT GGTTGTCT
4501 GCCTTCCCT TGTAAAGTC ACCTTGAAGT TAGTGCAGAA GAAATCAGAG CCCAGTCACA
4561 GAGTAATAT GGTCTGAAG ATTCCTTGT AGTGCCTCAGA ATCCATGACA TTTCAAGAGC
4621 CCTCTTGTG CCTTAAGTCA TTTGGGGTTG TATCTCTGC TTGATGTATG TGTGTGTT
4681 TATCAAAGAG TGAGATGGTT ACATAAGAGG TGCTCTAAG GACAGAGAGG ATTTGCAATT
4741 GTGGCATGTG ACATCCTCAG GCCTTGCCT GGTGCCAGGA GGAACGTGATG CAGAAAAGAG
4801 TAAGAGGTCA TTCTCTGGAG GCTGTCACTA TAGAGGAGAT CTTACAGTGC ATTCCCTCCT

Figure 23D:

4861 CCAGGCCCTG CCTGAGGATA GACATGTGCT GACTGCAACT GAAACAGAGG CTTGGGATGG
4921 AGAGTTAGGT TCACAGAAGG GAGGGTGGGA GATGGATGCT TGCTGGGTT C TGGGTCTCAT
4981 CACCAGCTCC TGACCACCCG GTCAGCCCAT GTGCTTATTTC CATA GCTTC TTTGCTATG
5041 TTTACTCAGT GTGGTGGTTG TTGGGACCCA GCAGAAGCCA GTCCCAGGCT GACAGCTGTG
5101 GATACACAGG GCAGCATGAG GGTCCCTCAGC CTGAAGCAGT CAGGCTGGCA GAAGAGAAAG
5161 ACCAGCACAC ATTCCCTCAA CCAACTATGT CTTGAAAAAC AAACATATTA TATCACATAT
5221 ATTGCAATT TGAGACAGCT AAAATGACT CGGGTAGCAT GACTCCAGGT GGGGATATCT
5281 GCAAGTGCCTA TGAGTGGCAG AGGGACAGCC AATGTGAGGC AAGAAGGAAT TCTGGCTCAA
5341 CACAGCTTAG CTCCCTGGTG TTGGTTCAAA CTTGAGAGT TTGACCACAA GCAC TTTATT
5401 TTTGACATAT TTAAACAGAG CACAACCTTG GAAAAAAGTT TTCTTATGAA AATTATCACA
5461 ATAAAGCTTA AGGCATGACT ACATTAAT GCCTTGCAA AGTATATGTG CCCTCTCCA
5521 CAAGAATGGT TCTATTGACT GAGAAATAAT GTTCAGGATA AAGATCCAGG AAGAAAAGAT
5581 CAGGGATAAAG TAAAATACTA AACTCTTGT CAAAGTACAT AGACCCCTTT TCATAACAT
5641 GGGTTCTATT GACTGACAAG CACTGCTCAG GAGTTGGGA AGAGTCTAGC ATAAGCACGA
5701 TAGCCTGGAG ACTCTAGTGA GGTCTAGTCT TACAGACAGC AAAAATCACC AGGTTACAAA
5761 CTACATTCCAT TTCCAGTTT CTGATCAGGC ACAGGTATGA ATCCCTCTG TTGAAGAGAA
5821 AAGTCCATGT GTTAAAATA TCTGGTTCT CCAGTGCTAT TAGCGAGAAG ACTTGAGCCC
5881 TATAACAACTC CCACCTGGAG TGACATCTG TCTTCATGGT ATATTACATA CCTAGACACG
5941 CTCATCTCAC AGACTTAGGA CTTTGTCTC TGATCTCCAT TTCTGATCCC ACTTCCACCT
6001 TTGCCTTGAT AGTGTCAATT TCTTCACTGC CTTGGTGACA ACCATGTTAT CCTCTGTGTA
6061 TTTGAGTGT ACCATTTCAT GATTTACCT GTATGCAAGA TCACACAGTC TTTGTCTTC
6121 TGTCTGGATG CATGCTAAC TCTACACAAAC AACCCCTCCC CGTCACTCAG ATCTTCTCC
6181 ATTAACACAT ACATGGTGCT GAAGAGGCTA GGGAGCTTC CTTCA GTGGGGAG GAGCTAGCTG
6241 GCTATTGGGC CTTTTGACT GTCCAGGAAG GCCCCCAATT GCTGAGACAA GAACTTAGAT
6301 TCTTCATTAT TGACTCTAAC TCATGTATCA AGCAGAACG AATGAATAGT TATCAACAGG
6361 ATCAGAGGTT CCAGTGTAAAG ACACCTTGAC ATGAAAGAAC GGAGGAAGGA CAGATGGATG
6421 CATAAAAGCA GGACCACTGC CCCAGGAAGG TCCTGGAAAC TGATGCAGGG CAAAGGACAG
6481 GTTATAAACC AAATCTTAGG GAGTCAGGAA GAGCACAGAG GAGCTCAACC AACTGACCAC
6541 TGCTTAGGGG CTACCAACCC AATCCTCCCT GTGGGAACAG CTAAGCTATC AGCCAAGGGT
6601 AATAAACAGG CAGGACCTGT GGATGACATG GAGAGCATAG GGACCCCTGGG TCCAGCCTTT
6661 AGCACCTGCA CTCTCAGGAT ACTCCACCAT TGTGTCTTAG AGAGCCTAGG GATACTGGGT
6721 CCAGCCTTG GTACCTTCAC TCTCAGGGTA CCCCCATCACT GTGTCTTGGA GAGCCTAGGC
6781 ACCCTGGGTC CAGCCTTCAG TACCTGCGCT CTCAGGACAC CCCACCAATT TCTCTTGC
6841 CGTCTCTTCTC TCCTCTTCTC CCCCCTCATT GTCTCTTCTC TGTGTTCTT C TTGACTCTCC
6901 TTTCCCTCA CACCCCTCACT CTAGTCTCC CTTCCCTCT CTGCATCACC CTATTCTCTC
6961 TGTGGTCCCT CCACCTTCCT TTATCTCTCA TGCTTCTCTC CTCCTCTCAA TACTTGTCA
7021 CCACTATACT TCAGGGGCA GCTCTAGTGA CAAAGCTGTT AATAGCAAGA CTCTCAGATC
7081 TCCAACGGCT CAGAGGAGCC AGACCCACCA AGAACTCTCT CCAGGTCCAA TTTCAGGTT
7141 CTTCGAAAGC TTTCAGCAAA TGCTCAGGGAA ACATGCCACT AACAAAGAAGA TGCAAATTCC
7201 AGTTGAGAGT GGGAAAGGCC CTTGCGTAGG TCCCCTCTC CAGGCCAAGG TCAGAGGGGC
7261 TCTGTGTAAT CCGGATTGAC AGGGCTCAGA ACAATGTTT GTTTTTAAGG TTTATTATT
7321 TTAGGTGTTA GTGTCTTGC TTGCATGACC TTATGTGCTAT CAGTGTGTG CAGGTTCTG
7381 ATGACAGTAG AGGAGGGCTT TGAATCCCTG GGGATAGGAA GTTACAGGAA ATTATAAGCT
7441 GCTTTGTGGG TCTTCTAGCT TTCCCAACAG AAGTGAATGC TCTTCACCAAC TGAGCCATCT
7501 CTCTAGGCC AAGAGACATT GCTTTATGGA TATAATTGTG TGTTGTGTC AACATTGAGG
7561 AAAGGGAAAT AAAAAAAA CTTCAGCCGC TAAGGTTGTA CAGTTCACT AATTGCTACT
7621 TTAGTTGTG ATAAAATGGC AGGTGCTTCA ACATTTATAT ATACAAAAC TTCCCTGCTG
7681 GTGGTTCAAC TGTGAGAAGT GGGGTAAGTG GGTGAGTTCT CTTTTCTGT CTCTGTCTCT
7741 GTCTCTCTCC TTCCATTCTT CTTAAAGGA AATAAACATT GCAGCTGGGT TATAGCTCAT
7801 CAATATGGAA GTACAGAAG TGAAAAAAGG CATTGCTTG GTGGGTGGTG TTACCAAGCTG
7861 ATTTTTGGTT GTCCCTGCAAG GAGGTCTGGG GACTGGCTGC TCTGTCTCTG TCTGTATGAG
7921 TGAGGGAAAGT CTGGGGAGCA GATTCCCTAA CCTTCAGGCTT GGCCTGGTTC CTGAGTGAAC
7981 CCAGCCTCTC TGGTCTAGT AGCTTTTCC AAACAGGAAT CTGAGTGGTG ACAGGAAACA
8041 AGTACCAAGCC CATTGCTTAA GTGCCAGGGT TAGTGAGGGC AGGAAGCTGC CATAGCTGGG
8101 ATTGAGTAGTT GTATTGGATG TAGGAAGTCC TATCCTGGGA CAGCTAATCC TTAATGCTTC
8161 ACTGGAGATT TTCAATGAGA AATTATCCC ACGGCCCAT A TGGCCCCATC CTTTGTCTC
8221 CAACAGCCAA GTATTTCCA TTAGAGGAGA CTTCCGTAC ACTTGATGGA TGCTCATTCC
8281 AAGGTGACTT GGGCGACTCA GTACAGACTT GGGATGACCT CTGACAGGCT AACCTCTCCC

Figure 23E:

8341 CAACAAGGGC CCTCTATGTT TGCTATGTA TGTAATGTCA GACATTGTCA GGAGGTCCG
8401 CAGCACAGCC TGCCCACTGT GAGGGCTCTC ATAGGTTTCC CACTGTCTTA TCTACACAGG
8461 GATAACGAGG AGGTAAGCTG CAGTTCCAG TCTCACTTCA CAGAGGAAGA GATAACCCCA
8521 TCCCAGGTCA TGTAGCCAGC AGTGGAAAGA ATGAGGATT GAACTCAGGT CTTCCAAGTC
8581 CCATTGATAG CATCTCCTCA CAAGTCCTT GCCACCCCTCA CGATGCCCTA GACACTTGC
8641 TGCCCTTAT ACTAAGGAGA TGCAGGTACA AGGGGTTAC CCATGTAGCA GCTGAGGCAG
8701 CTGGGGATAG ATACCAGCAG CAGGCCTGAT GTCACTTCA TAACCTCCAGC ATCCCCAGTC
8761 TGTGTTCTG GAGTGTGAAA ATCCCTACTT AACAAGATTG TGCAACAGTC CTTGGCTCTG
8821 TGACCCATAG CTGGAAACAG GATTCTCATT GATTTGTGGA ACATGGTGGC AGCCAGCCAA
8881 AAAGAGGGTC TGCATACAGA AGACACGTGT GGCAAGGCCA CAGCAGACTC TGACTACCTT
8941 AGCTTACAGA ATTACAAGGT CATAATGTCC TCTGCTTTGG TCACCTCATG TTAAGGACAG
9001 GCCCTAATGA AGATGGGGCA GAAGACTGAA GGAATGCCA ACCAATAACT GGCCCAACTT
9061 GAGACCCATC CTACAGGCAA GCATCAATT C TGACACTAC TAATGATACT CTGTTATGCT
9121 TGCAGACAGA AGCCTAGCAT AACTATCCTC CGAGAGGTCC ACCCAGCAAC TGACTGAAAC
9181 AGAAAAAGAT ATCCACAGGC AAACAGTGG A TGGAGGTCAG GGACTATTAT GGGAGAGCTG
9241 TGGGAAGGAT TAAAAAACCT GAAGGGATA GGAACCCAC AGGAAGACCA ACAGAGTCAA
9301 CTAAGAGACC TGTGGGAGCT CTCAGAGACT GAGCCACCAA CCAAAGAGCA TACACAGGCC
9361 GGTCCGAGGC ACCTGGCACG TGTGAAGCAG ACATGCAGCT CAGTCTCCAT GTAGGTCCCTC
9421 CAATAAGCGG TAGCCTGACT GCAGTATCCA ATCCCCAACCA GGGCTGCATA GTCTGGCTC
9481 AGTGGGGGAG GATGCCCTA ATCCTGCAGA GACTTGTGTA GTGGAGAGCT ATCCAGGGG
9541 AACCCACCCCT CTCTGAGAAC GGAATGGGG TGGGGGAGGG ACTCTGTGAA GAGGGACAA
9601 GGACAAACAA GAACCTAAA TAGGTCAAGGC CCTAAAGGCT TGCTAAGTAG CAGTGGCCCA
9661 GCTCTGTCT GTTCCTCAGC CCAAGGCTCA GCTCCCACCT GTTCTGTGT TTTCTGGCT
9721 TTTCATGGGC CTAGGACTTG GTGACCAGTT CAAACAATGG GGCCTGTGGA AGACACAATA
9781 TACAAGACTA GGGACATTCC TGTCTGCTG ACTATCCATA GCCTGATGTA GGTGGAAGGA
9841 CCCAATCACT GGATTCTAC CTTGCACAAA CCTTGACAGC TGAGGGCCTC TCAGAAACCT
9901 ATTTCTTCCA CTGAAAATG AGACTCTCAA ATGAACGTCG TGACAATCAT CAGGCTTATT
9961 AAAGAGGTGT ATCTAACCTG AATGGCAAGC AGACAGCAGG CAAATGTCTG TATCAACCTC
10021 TAGGAAGGAC AAGAACTGCT CACTGCTGCC CCCCAGGAGG CCATTGCTG AAACAGCTGC
10081 TCTCTGTCT GTGCACAGGC CCTGCCTCT CATTGCAGCC ACAGCCCCCTT CCTGTCTGAA
10141 CCTCTGTCA GGTCACTGGG AAACAGATCA AGATGGAACA GGACAGCTCC TGATGGTAAA
10201 TAAAAAACAG TGGTCATGGC TATTCA TAGG GTTTATGCT TCTTCAGTCC ACACTGTGAA
10261 GAGCTGTGGG CATGAACCCAC AGTGTTCGAG GTAGAGTTGG GTTCTGAAA TTCACAGTGG
10321 GGTGAGCTCA GTAAATGTGA GCTGGAGGTC ACTCGTGAGA CACACAGTCC TGCTGCTTCT
10381 GTTCCAATA TCCTGAGGAG ACGACACATC TACTTTGTC AGAGGCCACA GTCTAGTTGA
10441 CCTGAGAGTT ACCAGTTCT TATTGTGTG TGTGTGTG TGTTGTGTG TGTTGTGTG
10501 TGTGTTCTGT GTGTGAGTGC AGGTGCACAT ATGATAGCGT ACACGTTGAG GTCAAGGAG
10561 AACTATCAGG CGTTGTCCCC TCCTACTTT CCTCGGACTC TGGAGAACAA ACATGGGTCC
10621 TTATTCCAGG GGAGCAAGTC GCTGTTGGCT GACACATCTT GTCACATAC ATTTTACCTA
10681 GACAATGGAG CCTCCATCAG AGTATTACTT TAGCTCTCA CGATGGCAA TGCACCCACT
10741 CTCTACCCAC ATAGGAGTTG GGTCTCCACA CACCCACCA CCCCCTTCAC CAAAACGTTT
10801 TCAGTTACTT TATCTGGTAA AGTCATCAG AGAATGAAGC CAGTATTAAG AACATGGAAT
10861 CATTGGGAA CCTGGATCTA GCAATACCC ACCCTAGATG GAGTTGCTGA GTTTTCACCT
10921 CAGATTATAA TTCCCCCTA GCTTCTATGG TTTATTCTGA AACCAAGGGGA ACTCGATTCC
10981 TCCCCTTGGGAA CCACAGACAT CCTGGCTTGT GAATTACAT GTCATCTACT GCTAATCCAT
11041 TGGTAGTATG TGGCTCACAG AGACACACTA CAGTCATGGC CAATGTCAAG GTAGGACAGA
11101 TGTGAATCAT TCCCCCAGTC CTGCTGTTT CATGACTAAC CCTCCTCAGC ACAGTGACCA
11161 TGAACCTACT TTTCCCTC TTTTATTTT AGAATTGCTG GAATTCTGA TTTTGAGAAA
11221 TAATAGCCTT GGGCAGCATT AAACAAAATC ATCTAGAAAG CTGGTTTAAA ATACAGATGG
11281 TTGAGTCAGT GAAAGAGTGA GGAATGTCAT TATTGGCCCC TCACAGAGGC TGGCTCACTC
11341 CAGCAGAGGT GGTTGAAGCT CTTGGACACG GGTCAAGGTGC ATAGGAAAGG TNGTCTGGGA
11401 CACTGAGAAC CACAATTGAA CAAACAGAAC TGTTGGCTTT TTTTTTTTA AATGAGTTCT
11461 CAAAAAAATGA CTGGCTAGCT TAGGCAAATA CTTCGAGCCA ACCCAACAGA ACATCTTCC
11521 ATTGATTCTAT TCTGGATCTT CTTCTAGAC AATACTGAAC TGACCCCTTG TTGGCAGTCT
11581 CAAGTTTGAC AACATAGGGC TTTGAACCTG GCACAAGGTC CATCACTGTC ACCCAAGCCT
11641 CCTGGGTGAC CTTTGGGTTG GAATATCTTG GCTAACCTTA GATATTTCT TTGGAGTATC
11701 TTTAGAACAT CCAGGAATA GGGCTTGATT CTCATCCTGG GACCACAAATA TAAGTCACCC
11761 TAGAATCCCA GGAGATCGTG CAGAGAAACA AGGATCTCTC TCGTGTGCTT CTTCTTCATA

Figure 23F:

11821 AGCAGTGAGT AGTGACTCCA CTAAACTGAG TTCCCATCTG AGAGTCCACA GGAGGCTTTG
11881 GGGCAAGAAG CAGAGGGAAG GCACTTTTG TGTTGGTAAAG TTTTGACTC TAACAAATTT
11941 GAAGACATAG ATGACATTGT GTCAGACTAA CAACAACCTA GACTCATGTG GGTTCTGTT
12001 AGGGATCAGA TTTTATTCTAT CAATGACTTG TCTTAGTGTAGAGAAAGG CCTCCTACTG
12061 GAGTGTAGGC TCAATAATGAGA CAGAAGAGAT AGCTATTCTC CCTAGGGACT GTGCTGCTCC
12121 AAGTTTGGTG GAGAAAGGCA GTGGGGAAACC TAGATGTGCT CTCTGGGAG GGGGCTGAA
12181 GCTGGCTTCA TAGAAGGTGT GAAGTTTGC TGAAACATCT AACACAGAATT ATAGCTTAGG
12241 AAAGTGAGCA GGCAAGGCAG GGAATGTGTT GCATATGTAT ATGTACATGA ATATATTATG
12301 TTATAGATAC ACACACATT GAACCTCATT TGCAAGATGAC AGAAAATAGG TTATTTGCC
12361 TCTCTTAACG GCTAAGCACA ATGACTTCCA GTTCCATCCA TTTCCTGAAA TGCCACAATT
12421 TCATTTTCA TTGTTGGCTGA ATAAAATTCC ATTGCAGACT GGGCCCTACT TCATCCACTC
12481 CTGAGGGCAG GCATATCCCC TGGCTCATT TCTTACCTAT TGTGAAGAGA AGTGCAACTG
12541 TCTTGTGAA AGGCAAGCGT GAGAGAGGCA GGCACTAATT GTGGGTTTT GTTTCTCTT
12601 CCTGCTATGA CTCTCCATT GTCAAGAACCA AAGATCGATA AAAGCCGCCA CCATGAAAGC
12661 CATCTTAATC CCATTTTAT CTCTTCTGAT TCCGTTAACCC CCGCAATCTG CATTGCTCA
12721 GAGTGAAGCCG GAGCTGAAGC TGGAAAGTGT GGTGATTGTC AGTCGTATG GTGTGCGTGC
12781 TCCAACCAAG GCCACGCAAC TGATGCAGGA TGTCACCCCA GACGCATGGC CAACCTGGCC
12841 GTTAAACCTG GTTGGCTGA CACCGCGCGG TGGTGAGCTA ATCGCTATC TCGGACATTA
12901 CCAACGCCAG CGTCTGGTAG CCGACGGATT GCTGGCGAAA AAGGGCTGCC CGCAGTCTGG
12961 TCAGGTCGCG ATTATTGCTG ATGTCGACGA GCGTACCCGT AAAACAGGCG AAGCCTTCGC
13021 CGCCGGGCTG GCACCTGACT GTGCAATAAC CGTACATACC CAGGAGATA CGTCCAGTCC
13081 CGATCCGTTA TTTAATCCTC TAAAAACTGG CGTTTGCCTA CTGGATAACG CGAACGTGAC
13141 TGACGCGATC CTCAGCAGGG CAGGAGGGTC AATTGCTGAC TTTACCGGGC ATCGGCAAAC
13201 GGCCTTCGC GAACTGGAAC GGGTCTTAA TTTTCCGCAA TCAAACCTGT GCCTTAAACG
13261 TGAGAAACAG GACGAAAGCT GTTCATTAAC GCAGGCATTA CCATCGGAAC TCAAGGTGAG
13321 CGCCGACAAT GTCTCATTAA CCGGTGCGGT AAGCCTCGCA TCAATGCTGA CGGAGATATT
13381 TCTCCTGCAA CAAGCACAGG GAATGCCGGA GCCGGGGTGG GGAAGGATCA CCGATTACA
13441 CCAGTGGAAC ACCTTGCTAA GTTTGCATTA CGCGCAATT TATTGCTAC AACGCACGCC
13501 AGAGGTTGCC CGCAGCCGCG CCACCCCGTT ATTAGATTG ATCAAGACAG CGTTGACGCC
13561 CCATCCACCG CAAAAACAGG CGTATGGTGT GACATTACCC ACTTCAGTGC TGTATATCGC
13621 CGGACACGAT ACTAATCTGG CAAATCTCGG CGCGCACTG GAGCTCAACT GGACGCTTCC
13681 CGGTCAAGCCG GATAACACGC CGCCAGGTGG TGAACTGGTG TTTGAACGCT GGCCTGGCT
13741 AAGCGATAAC AGCCAGTGG A TTCAAGGTTTC GCTGGTCTTC CAGACTTTAC AGCAGATGCG
13801 TGATAAAACG CGCGTGTCA TAAATACGCC GCCCCGGAGAG GTGAAACTGA CCCTGGCAGG
13861 ATGTGAAGAG CGAAATGCGC AGGGCATGTG TTCGTTGGCA GGTTTACGC AAATCGTGA
13921 TGAAGCACGC ATACCCGCTT GCAGTTGTA AGGTACCCGG GGATCACAAAC TTGCCCCCTG
13981 AAGAGGAAGA ACAGAAGGAT GCCACAACTC TCCTGCTGGC TACTCTCCAG TGGTTTCATC
14041 TTACTTCTGA TGGCATTCTC CTCTAGAAAG TGCTACTATC ATCCACACAT TTCTACCTGA
14101 GACCACCCAA AGGACCTCC CAAATTCTCT TCCTCTCTGA GTAGTCTCCA CACCTGTTAC
14161 CACCATCCCA GAATTAAAAT CCTAACTGCA CTCTGGCGTG TGACTTGCCT CAGTCCTTGC
14221 AATAAGAGTT GTGGCAGTG CCAGGGTGG TGGCGCACCG CTTTAATTCC AGCACTGGG
14281 AGGCAGAGGC AGGGGGATTCT GTGAGTTCGA GGCCAGCCTG GTCTACAGAG TGAGTTCCAG
14341 GACAGCCAGG GCTATACAGA GAAACCTGT GTCGAAAAAC CAAAAAAA AAAAAAAAGTT
14401 GTTGGCAGAG TGTGGGTTAT ATACCAGGTG GAGATTCAA ATGAGTGGCT GAAGCTGTAG
14461 CCAGAAGGAA CTTAGAGGAT AGCTCATTAAC TAAAAAGAA ATGTAGAGAG TAGCAGAAC
14521 ATTGAGAGAG TGGCACACA GCCACTGTGT GAATGTGGCA GAACACAATC CAGCCAGCTA
14581 TACATGCATA AGTGTATATT GGCACCATCC TGACTGATGA GACACAGGAA AACAGATAGA
14641 CGGGGTTAGG TGGCCATGGC CTTTCTGCC TGCCCTCTCC TAAGGGTCAT CTCAGAACCT
14701 TATGCTCTCT TAACCTTCC ATTGCTACTT AGCTTCTAGA TATCACCTCC AGATTAGTCT
14761 CCTTGGGTAC ATCAGTGATC CTGGTGATAT CCAGGGCTTC CTGATTCCAT CTTTGTCTA
14821 GAGGCTGCAA CTAAGAGGTT CTTCTTAAAT CTTCACACCC TGATGCCAA AGGAAGACAC
14881 AGAAGTTCAC AGAGGTGAAG TGATTGATGT AGGACATACA GTGAGCAAGC ATCAGGGTCC
14941 GGATTATCTG ACTCTACTCT AACTTTATG TAAATGTGCT TTATGCCATT AACACTGTCA
15001 TTCTGTGCT TCAGCTCTGG GAGACTCCC AGCACTCTA GGCACAAAGCC ACAATTAAAGG
15061 GACTCTGACA CTCTGCATTG ATTAATTAGC ATGGTGGTCT CTATGTTTCC AGATTCTATG
15121 TTGTTTCACT TTCCATATAG GCTATGAGG GTGTGAGGAA ATTTTTGGG GACAGAATTG
15181 GAGGCAATCC ACCTCTCTCA GGAAGCCTCT ATCTGGAAAA GCTTACAAC CAGGGACAGT
15241 AACTGTAGGC CCAGTCCTTG GTGTCCAAAA TGGGTTTTAT GGTTGAATC TGCAAGCCT

Figure 23G:

15301 TCCATGTGCT CAAAGGTTG AACATGGAGC CTCCTCTGG TAACACTGTA TTGGAGGCTT
15361 TTGAGACTGG ATGCTCTTG GTCCCATGTT TTGCTACATC ATCTGTCAAG ATATGACCCA
15421 GGCATGCTAC CAGCTACAC AGACTATGCC TCTCCAGCTT TCATGTTCTC CCCACCATGA
15481 TAGACTTGTA TCTCCTAAAA ATGGAATCAA AGCAAACCTT TCCTGCATTA AGTTTTTTT
15541 TTCTGTAA GTGTTGGTC ACAGGGACAA GAAAACACTC AATACAGATA ATTAGTACCA
15601 GAGTTGAGGT TCATTGCTCT AGCAAGTGG ATCAAATTT TAGGGCTTTG GAACTGATT
15661 ATAAGAGACA TGTAGAAGAG TCTGAAGCTG TGGGCTACAG AAGTGTCAAC AGTTTTTAAG
15721 AATAGTTAA TACACCAGG GAATTGTGAA AATCAGAATG CTCACACAAA GGCAGACAGG
15781 AAAACGTGAG CATGTGGCGT GTGAGAGGGC ATAAGAAGGA ACCTAGGGGG AAATGAGCTA
15841 GAAGCCATTG GGCTACGTTA GGGAACGTGT GTGGCTGTG TTGGCCCAGC CCCTGGCAAT
15901 CTGAATGAGG CCAAATTTA AAGGAGTGG A CTAACTCGAT TGTCAGAGAA AATATCAAGA
15961 CAGACCACCA CTCAGGCTAT GCCGTGTTG TGACCGACCA GCTACTCTTA GCCAGCTCTA
16021 TTGTGAAATT CCAGAGCAAT TATCAGAGCA TGAAGATACA TACAGTTAG TGAAGTAAGG
16081 GGTGTGGGTC CCTAAGTGG A TGTTGCATAA ATCTATGTAG GTGATGCCA AGTACACATT
16141 GATAATCCAA AATATCAGCA ATGTGGATG TCTTCCAAGG AGACCTGTAG ACACACATT
16201 TAGAACTTTG CTCATGGCTG TAATAAATAG CTAGCTAGAA ATCATTTCTT GAAGAGGTTA
16261 GTCTGAGTTA CGGTTCCAGG GCAAACATTC AGTGTGGCA AGGAAGGCAT TGCAGTCAGG
16321 AGCCAAAGGT CAGCTGGTCA CATTGCACTA AGAGTAGAGA GTCAGAGTGT GAGTAGAAAG
16381 AGGATAACAGG TTATAAAACC TCACTGTCCA CTCTCAGCAA TCCATTTCTT CCTAAAAGGC
16441 TTTACCTTCTT AAAGATTTA GTCTTCAAA CCAGTACCG TAGCCTGGGA AAAAAAGTTG
16501 AAACAAATGA GCCTTGTGG GGCATTCAC ACTTAAACAA GGGCATCACC TAGGAGGAGC
16561 CCTGTGTGCA GTAGGAAGTG TGGCCTCTGT GTCAGGAATG CTCAGGCTAA TAAGGGGTCC
16621 TCTATCTGAG GGACCCTATG AAGATTCAAC AAGTAGTTGT GAGAATTCCC TGTAATGG
16681 TGCTACCAAT TTGACATTG TAGACCTGT ATTGTGTGCT TCTTTATTGG GCTCTCCCAT
16741 CTCCCAACTT TCCAACCCAT ATTCCACATT AATCCCTTCC ACCACCATGC AACACTAGGT
16801 AGGAGAGAAG GAAGGTTAGA AGAGAAAGTG GGTATAGATC TATTTAGACT ACTTCCGT
16861 GATTAGGGC AAGTCCAATC GTCATTGTCA GGATACCTCC AACCAAGCAAC CAGCAAACCA
16921 GCAAATCAGA AACAGCAAAA GCAGCCAACA AGGCAGCACT AACCAAGCAGG ATTGGGGTCG
16981 GTAGCGTGGG AGCAGTCACT ACTGGTCTTC TCATGGCTT GGCATTAATA CTCTCTCAAG
17041 AAATTCCGTA ATTTTTTCCC CACCACCTGA AATTCCGTA TTTTAAATGC AAACATATCTA
17101 CAGCTGGCAA AAATCACATC TCTCCTAGAG CACAAGACAA ATCATAGTTA CTGGCTATT
17161 GCAATCTGAA GCATCTCAAT ATCCCACACC TGGGATTAAAC AAAAAAACAT ATTCAACATCA
17221 CATAACTGTT TTTTTTTCC AATTTTTAT TAGGTATTCTT CTTTATTTCAC ATTCAAATG
17281 CTATCCGAA AGTCCCCAT ACCCTCCAC CTCCCTGCTC CCCTACACAC CCACCTCCAC
17341 TTTTGACCC TGGAGTTCCC CGGTACTGGG GCATATAAAG TTGCAAGAC CAAGGGGCCT
17401 CTCTCCCAG TGATGGCCGA CTAAGCCATC TTCTGCTACA TATGCAGATA GAGACACGAG
17461 CTCTGGGGGT ACTAGTTAGT TCATATTGTT GTTCCACCTA TAGGGTCGCA GACCCCTTC
17521 GCTCCTGGG TACTTTGTCT AGCTCCTCCA CTGGGGGCTC TGTGTTTAT CTAATAGATG
17581 ACTGTGAGCA TCCACTTCTG TATTTGACAG GCACTGGCCT AGCGTCACAT GAGCCAGCTA
17641 TATCAGGGTC CTTTCAGCAA AACCTTGCTG GCATGTGCAA TAGTGTCTGC GTTGGTGGT
17701 TGATTATGGG ATGGATCCAC TAGTTCTAGA GC

//

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER: _____**

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.